Intergenerational Impacts of Secondary Education: Experimental Evidence from Ghana*

Esther Duflo Pascaline Dupas Elizabeth Spelke Mark Walsh

August 25, 2025

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. Secondary school scholarships delay fertility, reduce unwanted pregnancies and change partner selection. Under-one mortality is reduced by half. We also document cognitive development gains of 0.24 standard deviations of test scores by age 5 among children of female scholarship recipients. Means-tested secondary school scholarships are a high return investment, with a conservatively estimated $4.1 in benefits for each $1 the policy would cost the government in the long-run.

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 (GHSERC12/04/17), and the IRBs of MIT (1406006442 and 1702862052), Stanford (40912) and IPA (77.08October-002). 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. We thank AER editor John Friedman and anonymous referees for helpful comments. The funding for this study was provided by the National Institutes of Health (NIH grant R01 HD039922), USAID-DIV (Award 7200AA20FA00002), the IGC, 3ie, the Partnership for Child Development, the Nike Foundation, the British Academy, the JPAL Post-Primary Education Initiative and the Marshall Foundation. We thank them, without implicating them, for making this study possible. This manuscript is subject to the NIH Public Access Policy. Through acceptance of federal funding, NIH has been given a right to make this manuscript publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH.
Duflo: MIT and Collège de France (eduflo@mit.edu); Dupas: Princeton University (pdupas@princeton.edu); Spelke: Harvard University (spelke@wjh.harvard.edu); Walsh: GiveWell.

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. (2025), 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,

[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

especially women, may choose and be able to have fewer children, delay childbearing, and invest 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. (2025), 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 ("first generation") 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 ("second generation") 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. The tests were 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 receiving a scholarship impacts family formation. At our first follow-up in 2013, female scholarship recipients were 7 percentage points (14%) less likely to have had a pregnancy. This effect was driven by a 7 percentage points (17%) decrease in unwanted pregnancies. By 2016, they were 12 percentage points (24%) less likely to have ever lived with a partner. By 2022, their number of children had mostly caught up (1.8 children among non-recipients and an insignificant point estimate of -0.135 for the effect of the scholarship); but the catch-up was very gradual. These outcomes are harder to observe and measure for men. Yet, the effects for male scholarship recipients, while less precisely estimated, are remarkably similar in percentage terms. Where the results differ starkly between women and men is on partner selection. Among those who had started cohabitating by 2019, the scholarship treatment effect on the partner's level of education is *positive* for women and *negative* for men.

The second result is that children of 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 4.2% before the age of three. Survival increased by 1.7 and 2.0 percentage points (p-values 0.038 and 0.042), respectively, among children of female scholarship recipients. This represents a halving of under-one and under-three second-generation mortality in our sample. The sample of second-generation children among males in the study is small, and likely negatively selected, because men start having children later than women. This is especially true for outcomes that we can only measure for children born three or more years prior to the end of our study period. We observe a 1.4 percentage points decline in under-1 mortality among children of male scholarship recipients, but this is entirely driven by children born in the latter half of the sample period. For children born before 2017 to men in our sample, the point estimates on under-1 and under-3 mortality are (insignificantly) *negative*, consistent with a negative selection into early fertility for the treatment group.

The third set of results concerns cognitive development for second-generation children. To avoid bias stemming from the fact that individuals who received a scholarship started having children later in life, which implies that their children tend to be younger, we collected data on second-generation 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 second-generation 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

0.24 standard deviation higher (p-value 0.005), and by seven it is 0.25 standard deviation higher (p-value 0.035) if their mother received a scholarship. Those are large impacts, found for most of the cognitive domains tested. In contrast, we find no evidence of positive cognitive development impacts for children of male scholarship recipients. The point estimates for the treatment effects are in fact negative for children of male recipients at 5 and 7 years—measurements that could only be done among children born early in the study period.

An important mediator of the impact of the scholarship program on second-generation outcomes seems to be the education of the child's primary caregiver. In most instances, the primary caregiver is the child's mother. This means that children of female scholarship recipients have a much higher chance of having an educated caregiver; but this is not true for children of male scholarship recipients. In fact, male scholarship recipients who start cohabitating early tend to choose partners who are relatively less educated than partners of male respondents in the control group. The opposite is true for partnerships that form later. This could explain why we see consistently large positive impacts on children of female scholarship recipients, but heterogeneous impacts on the children of male scholarship recipients.

In turn, the education of the primary caregiver (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 (including self-employment earnings) in the first 12 years after the scholarships were issued (Duflo et al., 2025).[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 for the second generation (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 stimulation of children. Children of mothers who received the scholarship have received more preventive care, and mothers who received a scholarship are more likely to report counting, naming or drawing things with their children, singing them songs, etc. 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

[3] A positive effect on the earnings of female scholarship recipients does emerge after 12 years (Duflo et al., 2025), driven by relatively few women who obtain a job in the government sector.

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 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, possibly owing to the fact that children of male scholarship recipients do not have a more educate female caregiver, and are less likely to live with their father.

Benefit-cost analyses valuing a life at 100 times the Ghanaian GNI per capita (the medium scenario recommended in the literature) suggest that means-tested secondary school scholarships are a highly valuable investment in contexts similar to Ghana's, with benefits at least three times greater than the costs. The estimated marginal value of public funds is 4.1 when both men and women are eligible, and 4.6 when scholarships target women only.

This paper makes a number of contribution. First, it 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 delays fertility. An experiment in Malawi (Baird et al., 2019) finds that conditional cash transfers increase secondary education and reduce fertility among adolescent females. 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. 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) find 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. Exploiting variation in college costs during a mother's adolescence, Carneiro et al. (2013) find important intergenerational effects of maternal education on children's cognitive development 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).

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 in a low- or middle-income country.[4] 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 and government education expenditures with self-reported school expenses and wages during secondary school, we accurately estimate the total social cost of secondary school scholarships, an important component for cost-benefit analyses.

[4] Walker et al. (2023) estimates intergenerational child mortality impacts of a randomized *health* intervention—school-based deworming—in Kenya, and find a 22% reduction in under-5 mortality among offspring of individuals exposed to the program in their primary school years. The mechanisms of impact appear somewhat similar to those we identify, since not only the health but also the education of the first generation improved.

2 Setting, Experimental Design and Data

In 2008, Duflo et al. (2025) 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. (2025)). 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. (2025) 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).

The scholarship study team called the scholarship-lottery participants once a year from

[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.

2009 to 2012, to update their contact information and measure key outcomes (education status, family formation). In 2013, a detailed in-person follow-up survey measured schooling, occupation, cognitive skills, labor market expectations, 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. (2025) report 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). The effect of scholarships on SHS completion is large and statistically significant at conventional level at all quartiles of the initial test score distribution. In 2013, scholarship recipients scored 0.16 standard deviations higher on 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 this difference in treatment effects between gender groups is not statistically significant (Duflo et al., 2025, Table 1).

Overall, as of 2022, the scholarship had led to an average increase in total years of education of 1.46 years for women and 1.18 for men. While this increase is mainly due to more years of secondary education, there were 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 4.0 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 (Duflo et al., 2025, Table 2).

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) (Duflo et al., 2025, Table 3). 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 2023 (Duflo et al., 2025, Figure 2).

The education and labor market effects of the scholarship are nearly identical for the subsample whose offspring form this paper's sample for the analysis of intergenerational cognitive development effects, with one exception (Table A1). Within this subgroup, there is already a significant gap in earnings between women who received the scholarship and those who did not as of 2019.

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

[6] In their 2017 *Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries*, Fernald et al. (2017) review the weaknesses as well as the strengths of existing child measures, and highlight that direct testing of infants (0-18 months) shows little correlation with later child outcomes (p.31). 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).

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).

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]

[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.

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 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 A1 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,

[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 months old using the 60-69 month old instrument. The surveyors were permitted to survey children up to 99 months old using the 84-96 month old instrument and up to 39 months old using the 30-36 month old instrument. We control for child age in months in the analysis.

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., 2025).

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. (2025).

3.2 Second-generation: Child survival

To study impacts on child survival among the second generation, our sample consists of children of scholarship-lottery participants. Mechanically given the impact of the scholarship on fertility,

[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.

children of female scholarship recipients are younger than children of females in the control group (4.71 months younger; p=.067; Table A4 row 1). We therefore focus on survival to age 1, 3 and 5 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, or 5) by the time their scholarship-lottery participant parent was last surveyed. We run the following regression:

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

where Y_{ij}^{a} is the outcome (*child survived to age* a) for child i of scholarship-lottery participant j; T_j is an indicator that the scholarship-lottery participant j 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 row 2), 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

[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.

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 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 row 6), 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.

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 treatment and control groups (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 child survival, 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

[13] For reference, Figure A2 plots cognitive development score against child age in months by age window among the control group.

(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. This highlights the benefit of having conducted surveys almost yearly. Drawing upon all follow-up surveys from 2009-2023, we have survival status for 93% (98%) of all children ever mentioned by a female (male) scholarship-lottery participant and who had, or would have, turned at least 1 by the last follow-up survey in Spring 2023 (Table A2).

A limitation of our child survival data is that we miss children who would have turned 1 or 3 or 5 after we stopped surveying in Spring 2023. Of the children ever mentioned by a scholarship-lottery participant, 8% had not turned 1, 24% had not turned 3 and 41% had not turned 5 by Spring 2023. There may also be a sample selection bias among children ever born. Since Ghana has a total fertility rate of ≈3.4 per woman, it is likely that our sample will have more children over the course of their lives.[14] 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 (Table A3, Panel A). In particular, among men who had a child by Spring 2023, scholarship recipients are 9.7 percentage points (+25%; p-value 0.055) more likely to have grown up in a household with no male head 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).[15]

For the cognitive tests, we had relatively high tracking rates. Tracking rates ranged from 78-96% for eligible children (Table A2). 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).[16] Overall, we have measures of cognitive development for at least one eligible child for 64% of female

[14] Total fertility rate estimates for 2023 as reported in the World Bank Gender Data portal: `https://genderdata.worldbank.org/en/indicator/sp-dyn-tfrt-in`, last accessed 6/30/2025.

[15] 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.

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

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 those with children in the cognitive games sample have slightly different characteristics between treatment and control groups (Table A3, Panel B.)

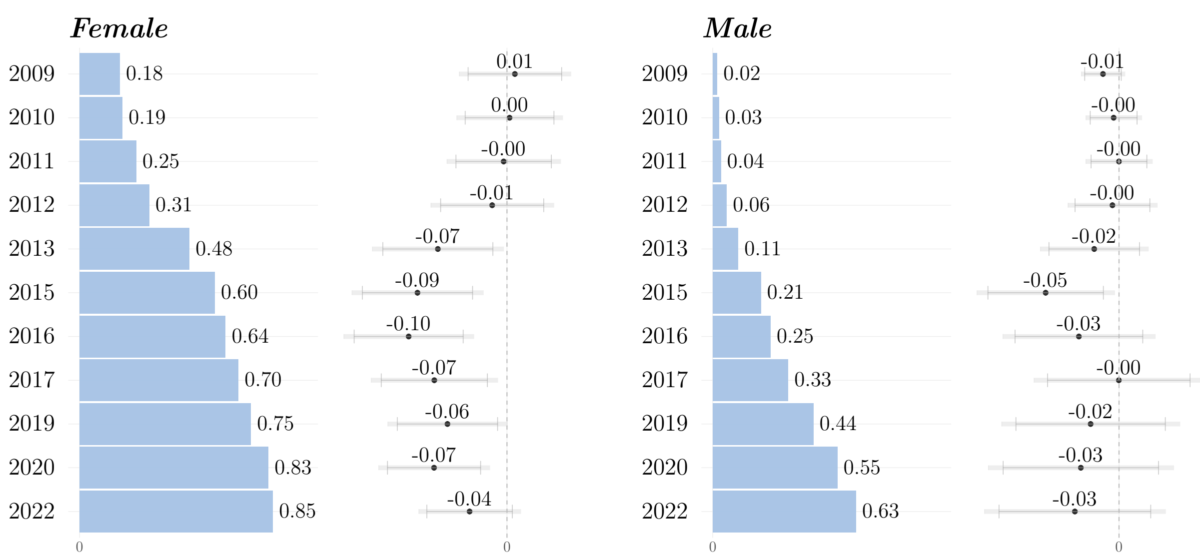
4 Results

We start by presenting the impact of the scholarship on family formation and fertility choices. We then show second-generation impacts on survival and cognitive development. Finally, we present evidence on possible mechanisms for the second-generation effects, including parent/caregiver characteristics and behavior.

4.1 First-generation: Fertility and family formation

Figure 1 and Table 1 present the impacts of the scholarship on fertility and marriage. We discuss the impact on women and men in turn.

Figure 1: Impact on childbearing onset—ever pregnant/ever had a pregnant partner



Notes: Notes: Data from 2013 in-person follow-up and yearly phone surveys. Survey year indicated on the left. The outcome shown is "Ever pregnant" (for females) and "Ever had a pregnant partner" (for males). Left half of each panel shows means in comparison group; right half shows estimated treatment effects of scholarship. 90% CIs at whiskers and 95% CIs at the end of the confidence bars.

Female scholarship recipients By 2013, women in the scholarship arm were 7 percentage points less likely (p-value 0.039) to have ever been pregnant (on a base of 48% 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 (Table 1, column 2). 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 (-0.15 (11%) fewer children, p-value 0.065). 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 appears to partially catch up for female scholarship recipients. By 2022, women are 4 percentage points less likely to have ever been pregnant, and the number of children ever had is 7% smaller, but these differences are 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 "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. (2025), 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).

Table 1 documents 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 percentage points (24% of the control mean) less likely to report having ever lived with a partner. As of 2019, they are 6 percentage points (p-value 0.067) less likely to be married or cohabiting (a 13% drop from a base of 47.5% in the control group; see Figure A3a for the effects year by year). Conditional on having a partner, they are 7 percentage points (36%) more likely to have a partner who completed tertiary education (p-value 0.071). The way in which they met their partner has changed, with a significant increase in the role of school and friends and a decrease in workplace matches (Figure A3b). There is however no significant change in the age gap with the partner (Figure A3c).

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
Stepdown P-value	0.188	0.188	0.243	0.300	0.003	0.243	0.931	0.243
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
Stepdown P-value	0.761	0.824	0.857	0.857	0.124	0.373	0.071	0.117
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

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. The last row shows the p-values for tests that the effects are identical between males and females. 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-values; the fourth row reports the family-wise Romano-Wolf step-down p-values; comparison group means are in the fifth row; the sixth row reports the number of observations. Controls include JHS finishing exam score and baseline region fixed effects. *** p<0.01, ** p<0.05, * p<0.1

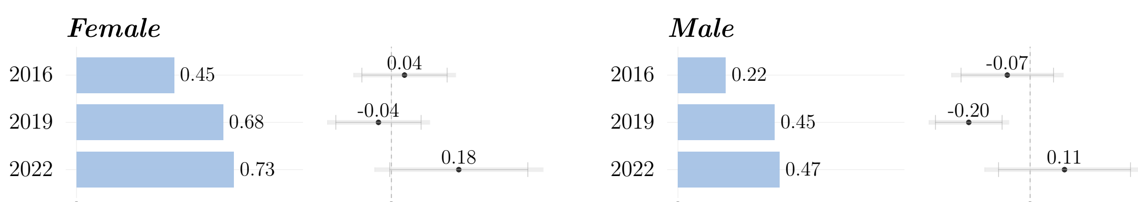
In the rest of the paper, we show evidence that is consistent with either channel (4) (reduced costs of investing in children quality) or a direct impact of the "wantedness" of children on their quality. Indeed, children of female scholarship recipients are healthier, and they have higher cognitive achievement.

Male scholarship recipients Much like women, young men who were awarded the scholarship delay family formation. The point estimates for fertility are smaller and not statistically significant, although it is worth nothing that they are similar in percentage terms as those found for women, because control men report having fewer children. We see a -1.8/11.2=16% decline in ever having a pregnant partner (column 1 Panel B of Table 1, p-value 0.37)

One important measurement issue is that parenthood is likely measured with much more error for men than for women. Since many pregnancies are out of wedlock and not all of them lead to marriages, it is possible that male respondents under-report (or are unaware of) births they may have been responsible for. It is also possible that this is differential by treatment status. Some men with scholarship could be more likely to acknowledge paternity (because they are more desirable marriage prospect for example), which would lead us to underestimate the treatment effect. Some may be less likely, if they are delaying family formation because they are hoping to attend tertiary school. This would lead us to overestimate the treatment effect.

Turning to cohabitation, the magnitude of effects in percentage terms is very similar to that for women, even though men start cohabitating at a later age. We see a 25% lower chance of ever having had a partner as of 2016, and a 16% lower chance of cohabitation in 2019 (p-values 0.027 and 0.117). Unlike women, however, male scholarship recipients are also significantly more likely to live with their parents (+ 8 percentage points, or 30% of the control mean, in 2019; p-value 0.011). Among those who had a partner by 2019, the partner is significantly *less* likely to have attended tertiary education, which is the opposite of the pattern we see for women. This may reflect a negative selection of the men who start cohabitating early, which, as we discussed in section 3.4, we also see among men who have children.

Figure 2: If started cohabitating in the last three years: Partner has completed SHS



Notes: Observations are female (left panel) and male (right panel) scholarship-lottery participants. The 2016 estimates include all respondents ever cohabiting by 2016. For the 2019 and 2022 estimates, we focus on the subsample who report cohabitating with a partner for the first time in the previous three years. Bars indicate the comparison group mean. Black dots are point estimates of the treatment effects, with 90% and 95% confidence intervals indicated by the line and whiskers, respectively. The share of study participants who report ever cohabitating by survey year is shown in Figure A3a.

Figure 2 shows how the characteristics of partners change across survey years. Male scholarship recipients who started cohabitating early have relatively less educated partners, while those who initiate cohabitation later have relatively more educated partners, with a gap of 31 percentage points in the treatment effect on the likelihood that the partner has completed SHS, between those who started cohabitating in the three years leading up to 2022 (+0.11) vs. the three years leading up to 2019 (-0.20). We come back to this heterogeneity when we look at outcomes of offspring in the next section.

4.2 Second generation: survival

Table 2 presents the results on child survival. The unit of observation in this table is the child of a scholarship-lottery participant. We show two specifications per panel: the basic OLS, for which we adjust for multiple hypothesis testing using the step-down procedure proposed by Romano and Wolf (2005); and the specification with entropy balancing, which reweighs observations so as to obtain balance on 2008 characteristics and thus helps understand whether the estimated effects are driven by differential selection into family formation by treatment status.[17]

Among children of female scholarship-lottery participants, we find a large, significant, and very consistent reduction in mortality in the treatment group. Specifically, we find a 1.7 percentage point increase in the probability of surviving until age 1 (49% decrease in mortality, p-value 0.038, step-down p-value 0.005), an increase of 2.0 percentage points in survival to age 3 (a 48% decrease in mortality, p-value 0.042, step-down p-value 0.006), and an increase of 2.4 percentage points in survival to age 5 (a 43% decrease in mortality, p-value 0.077, step-down p-value 0.013). The results are unchanged with entropy balancing—as expected since there is no imbalance between the recipients who had children and non-recipients who had children among women (Table A3). Results also barely change when we replace birth year fixed effects with a linear control in birth year, drop birth order fixed effects, add a control for mother's age at birth, or drop all controls for birth year (Table A6).

For children of male scholarship-lottery participants, the sample sizes are considerably smaller, so the estimates are noisier. What is more, since men have children later than women, the sample for which we have survival to age 3 (even more so for survival to age 5) is

[17] The entropy balancing method of Källberg and Waernbaum (2023) 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, total number of household members, highest education of household head: SHS or more, highest education of household head: primary or less.

Table 2: Second Generation Impact: Child Survival

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All children			Born 2009-2016			Born 2017-2022	
Survival to:	1 year	3 years	5 years	1 year	3 years	5 years	1 year	3 years
Panel A: Children of female scholarship-lottery participants								
Treatment	0.017**	0.020**	0.024*	0.015	0.023*	0.020	0.024**	0.020
	(0.008)	(0.010)	(0.013)	(0.012)	(0.013)	(0.014)	(0.012)	(0.016)
P-value	0.038	0.042	0.077	0.218	0.069	0.155	0.040	0.223
Stepdown P-value	0.065	0.065	0.078	0.230	0.111	0.218	0.063	0.232
Entropy balanced								
Treatment	0.017**	0.020**	0.024*	0.015	0.024*	0.021	0.024**	0.023
	(0.008)	(0.010)	(0.013)	(0.012)	(0.012)	(0.013)	(0.012)	(0.016)
P-value	0.038	0.036	0.073	0.199	0.051	0.115	0.043	0.147
Comparison mean	0.965	0.958	0.947	0.967	0.958	0.956	0.963	0.959
N	1,707	1,395	1,084	947	935	899	760	460
Panel B: Children of male scholarship-lottery participants								
Treatment	0.014	0.008	0.000	-0.006	-0.007	-0.015	0.029**	0.030*
	(0.010)	(0.012)	(0.018)	(0.017)	(0.017)	(0.020)	(0.012)	(0.018)
P-value	0.149	0.479	0.985	0.707	0.708	0.452	0.014	0.096
Stepdown P-value	0.269	0.621	0.987	0.739	0.739	0.585	0.029	0.090
Entropy balanced								
Treatment	0.016*	0.013	0.008	-0.004	-0.004	-0.012	0.032***	0.038**
	(0.009)	(0.011)	(0.017)	(0.015)	(0.015)	(0.018)	(0.010)	(0.016)
P-value	0.068	0.242	0.633	0.812	0.804	0.520	0.003	0.014
Comparison mean	0.970	0.971	0.968	0.979	0.978	0.977	0.964	0.962
N	1,016	772	520	420	411	390	596	361
P-val male=fem	0.631	0.405	0.284	0.290	0.159	0.159	0.923	0.655

An observation is a child of a participant in the lottery for secondary school scholarships. "Treatment" means the child's parent won the scholarship lottery for Senior High School (SHS). Panel A shows results for children of female scholarship lottery participants; Panel B shows results for children of male scholarship-lottery participants. Columns 1, 4 and 7 (respectively, columns 2, 5 and 8 and columns 3 and 6) include all children who had reached 12 (respectively, 36 and 60) months as of the 2023 survey. 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-values; the fourth row reports the family-wise Romano-Wolf step-down p-values; comparison group means are in the fifth row; the sixth row reports the number of observations. Controls include year of birth fixed effects, birth order, region fixed effects and the JHS finishing exam score of the scholarship-lottery participant. *** p<0.01, ** p<0.05, * p<0.1.

differentially selected, and comparing across columns (e.g., survival to age 5 vs. survival to age 1) does not tell us whether a given child survived to these age milestones, as the columns consider different *samples* of children. For this reason, in columns 4-8 we explore heterogeneity by child birth cohort, to parse out composition from treatment effects.

Turning first to average effects, we estimate an insignificant 1.4 percentage point increase in the probability to survive until age 1 (47% decrease in mortality, p-value 0.149, step-down p-value 0.269) and an insignificant increase of 0.8 percentage points in survival to age 3 (a 31% decrease in mortality, p-value 0.479). While smaller, these estimates are not significantly different from the estimates for female respondents. With entropy balancing, the reduction in infant mortality before age 1 becomes significant at the 10% level for children of male scholarship recipients.

Pooling children of female and male scholarship-lottery participants, the treatment effects on survival to age 1, 3 and 5 are 1.7, 1.4 and 1.5 percentage points, with p-values of 0.005, 0.063 and 0.158 and step-down p-values of 0.022, 0.045 and 0.130 (Table A5).

Heterogeneity across birth cohorts Turning to the split at the median year of birth (columns 4-8 of Table 2), we find positive survival impacts for both cohorts among children of female scholarship recipients. In contrast, we observe some important heterogeneity for children of male scholarship recipients: the survival effects are completely absent among children born "early" (i.e., before 2017) to male scholarship recipients, but they are quite large and significant for children born 2017 or later (+2.9, p-value 0.014, step-down p-value 0.029). This heterogeneity could reflect differential selection: treatment men who had children relatively early may have been negatively selected. The fact that the point estimates for the earlier birth cohorts become less negative with entropy balancing partially confirms this, though they remain negative. We discuss the role of negative selection of *partners* in section 5.

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. 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, 50/1,707 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 Second generation: cognitive development

Table 3 shows the results on child cognitive development; once again, the unit of observation is the second-generation child.

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
Stepdown P-value	0.868	0.929	0.929	0.027	0.141
Entropy balanced					
Treatment	-0.025	-0.060	0.011	0.235***	0.247**
	(0.082)	(0.112)	(0.080)	(0.081)	(0.107)
P-value	0.758	0.591	0.891	0.004	0.022
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
Stepdown P-value	0.577	0.505	0.970	0.310	0.856
Entropy balanced					
Treatment	0.109	-0.220	-0.055	-0.263**	-0.211
	(0.110)	(0.149)	(0.086)	(0.118)	(0.151)
P-value	0.323	0.142	0.528	0.026	0.163
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

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). See Table 2 notes on how to read the table. All regressions control for child age in months, child gender, child birth order, measurement round fixed effects, and fixed effects for scholarship-lottery participant baseline region and JHS finishing exam score. 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 and 7 years because these tests were introduced later (July 2021 and May 2019, respectively). *** p<0.01, ** p<0.05, * p<0.1.

For children of female scholarship recipients, the estimated treatment effects are insignificant and slightly negative for 18 month olds and 2.5 year olds, and insignificant and slightly positive for three year olds. In contrast, the five and seven year olds of female scholarship recipients score substantially higher on the cognitive development tests, 0.24 and 0.25 SDs respectively (p-values 0.005 and 0.035). These results are robust to multiple hypotheses testing adjustment and entropy balancing. They are also robust to excluding measurement round fixed effects (Table A7).

The effects for children of female scholarship recipients at ages 5 and 7 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.[18,19]

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 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.[20] It is thus possible that the null effects on younger age

[18]Note that Coffey and Spelke (2023) document 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.

[19]We break down the impacts on child cognitive development by child gender in Table A8. 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.

[20]If the tests are measuring a cognitive domain accurately, performance on one of the tests should predict

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 A9).[21] 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.[22]

Turning to children of male scholarship-lottery participants, we find no significantly positive impacts on cognitive development at any age, and the point estimates are in fact 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, step-down p-value 0.104). 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). Note that almost all children who could be tested at 5 or 7 years old were born before 2017, so, for children born to male scholarship recipients, the non-positive effect on cognitive scores is consistent with the non-positive effect on survival observed in Table 2 columns 4-6.

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

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 A4.)

[21] Dupas et al. (2025) 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.

[22] One could also think that the treatment effect is observed at older ages because of a *selection effect*: due to delay in the onset of childbearing for the treatment group, children eligible for the 5- or 7-year tests include, in the control group, *children who would not have been born* had their mother received a scholarship. This cannot explain the results, however because we did not start testing children for cognitive development until 2017, which means that we missed the "window" to administer the 5 (7)-y.o test to children born before 2012 (2010). Thus, even though women in the treatment group had 0.11 fewer children aged at least 7 by 2023 compared to women in the control group (p-value 0.06), there is no significant difference between treatment and control arms in the number of children administered the age-7 cognitive test. The same goes for age 5 (Results available upon request). Thus these effects are by construction coming from the "intensive" margin of a similar sample of children born in treatment and control group, doing better in the treatment group.

(Table A3). However, the results on offspring cognitive scores are qualitatively unchanged when we use entropy balancing, suggestive that the results are not due to negative selection.

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 impacts.

5 Mechanisms

What are the likely explanations for the positive impacts for children of female scholarship recipients, and the lack of effects for children of male recipients born before 2017?

5.1 Parental education

Maternal and paternal 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 (Table A10). As a result, the primary caregiver for the children of female scholarship 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 1 and 2). In contrast, children of male scholarship recipients have caregivers with the same level of education as children of males in the control group.

While the caregivers are almost exclusively females, the partners of caregivers are almost exclusively males. Mechanically, there is a large treatment effect on the secondary school status of partner of caregivers for children of male scholarship recipients (i.e., the partner of the caregiver is typically the scholarship recipient father). There is also a positive (and quite large in percentage term) effect on the education levels of partners of caregivers of children of female scholarship recipients, though not significant (columns 3 and 4 of Table 4).[23] Controlling for partner's education does not change the treatment effects for children of female scholarship recipients however, suggesting that maternal education, not partner education, is the main driver (Table A11).

Another reason why paternal and maternal education may not have symmetric impacts on a child's environment is that children of male scholarship recipients appear significantly less

[23]The effect on the education of partners of caregivers is smaller than the effect on partners of scholarship recipients in Table 1, likely due to the fact that the sample in Table 4 is selected on children being old enough to be tested, and this likely includes scholarship recipients most likely to have found a partner while enrolled in tertiary.

likely to live with their father (-8.7 percentage points, p-value 0.024, column 4 of Table A10). 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%).

Table 4: Caregiver Characteristics

	(1) Caregiver completed secondary education	(2) Caregiver completed tertiary education	(3) Caregiver's partner completed secondary	(4) Caregiver's partner completed tertiary	(5) Earns income	(6) Public sector job	(7) SES index
Panel A: Children of female scholarship-lottery participants							
Treatment	0.248***	0.053**	0.063	0.040	0.021	0.052**	0.067
	(0.040)	(0.021)	(0.049)	(0.037)	(0.030)	(0.022)	(0.063)
P-value	0.000	0.010	0.202	0.274	0.491	0.017	0.289
Stepdown P-value	0.000	0.074	0.586	0.631	0.631	0.100	0.631
Comparison mean	0.218	0.035	0.432	0.155	0.748	0.039	-0.005
N	2,756	2,756	2,308	2,308	2,756	2,756	2,718
Panel B: Children of male scholarship-lottery participants							
Treatment	0.010	-0.002	0.204***	0.025	-0.049	-0.027***	0.040
	(0.036)	(0.013)	(0.055)	(0.029)	(0.034)	(0.010)	(0.075)
P-value	0.785	0.858	0.000	0.400	0.153	0.009	0.596
Stepdown P-value	0.951	0.951	0.011	0.871	0.583	0.087	0.933
Comparison mean	0.196	0.026	0.453	0.070	0.818	0.033	0.002
N	1,536	1,536	1,341	1,341	1,536	1,536	1,516
P-val male=fem	0.000	0.026	0.064	0.617	0.140	0.002	0.648

*** 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). All regressions control for child age in months, child gender, baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. Components of the SES index are shown in Table B1.

To understand the heterogeneity in treatment effects on survival across birth cohorts among children of male scholarship lottery participants, we cannot rely on the caregiver survey data, which is, by construction, only available if a child survived to age 1.5 and was included in the cognitive test sample. Instead, we use data on scholarship lottery participants and their partners. Here again, maternal education appears to be the key factor. Indeed, recall from Figure 2 that treatment effects on partner selection changed over time for male scholarship recipients, going from negative to positive. Assuming that, for the majority of cases, the partner of the father is/was the mother/caregiver, this would mean that the children born early to male scholarship recipients have mothers with somewhat lower education (which could explain the negative point estimate on survival if maternal education is the key factor), while those born later have more educated mothers (consistent with the positive point estimate on survival if maternal education is the key factor).[24]

5.2 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 0.35 years for children born to female scholarship recipients (p-value 0.147) compared to the control group (column 7 of Table A6). For first-born children, the gap in maternal age at birth between treatment and control group rises to 0.55 years (p-value 0.074, Table A12, 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 A12 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 A6, -0.46 years; p-value 0.22), yet the point estimate on survival is positive.

Resources Duflo et al. (2025) report no significant 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

[24] Education and other characteristics of scholarship-lottery participants and their partners, and associated treatment effects, broken down by child birth cohort for the survival-to-1 sample, are shown in Figure A5.

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 that children of female scholarship recipients are more likely to have a caregiver who has a public sector job (column 6 of Table 4).[25] Yet, we see no difference in the pre-specified SES index that includes information on per adult equivalent food spending and dwelling characteristics, column 7 of Table 4). It could be that our SES index misses some the resources available, such as savings for emergencies.

Resources *per child* could also 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 7 of Table A10). 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 7 of Table A10 is found even among biological children.

Caregiver practices, behaviors and beliefs Table 5 turns to parental/caregiver behavior, and a number of meaningful differences emerge. Turning first to health investments, we find that children of female scholarship recipients are significantly more likely to receive prenatal care (column 1)—prenatal care is quasi-universal, so the 1 percentage point increase in prenatal care for the treatment group corresponds to almost a halving of the odds of *not* having received prenatal care. We see a positive and significant (p-value 0.068) effect on preventive health behaviors (column 2).[26] Moreover, we observe an improvement in caregiver-reported child

[25] For the sample of children for whom we have information on survival, we also see that mothers in the treatment group are more likely to have a partner with a public sector job compared to mothers in the control group (Figure A5).

[26] The preventive health behaviors in the index are shown in Table B2. We see significant increases in access to bednets and private toilets, 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).

health (Table A13 and Table B3). Finally, for the subset of children for which anthropometric outcomes could be measured (see breakdown in Table B4), 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.

Table 5: Parental / Caregiver Behavior, Aspirations and Beliefs

	(1) Received prenatal care	(2) Preventive health behaviors index	(3) Child investment index	(4) Schooling Index	(5) Child stimulation index	(6) Beliefs: Role of parental stimulation index	(7) Aspiration: child's years of education
Panel A: Children of female scholarship-lottery participants							
Treatment	0.010**	0.120*	0.010	0.020	0.147**	0.046	0.025
	(0.004)	(0.066)	(0.053)	(0.069)	(0.059)	(0.075)	(0.039)
P-value	0.024	0.068	0.850	0.770	0.013	0.546	0.527
Stepdown P-value	0.160	0.300	0.948	0.948	0.104	0.944	0.944
Comparison mean	0.989	0.014	0.047	0.065	-0.006	0.060	16.753
N	2,977	2,743	2,742	1,833	2,739	2,741	2,726
Panel B: Children of male scholarship-lottery participants							
Treatment	0.007	0.006	-0.047	0.062	-0.097	0.001	0.098
	(0.009)	(0.079)	(0.074)	(0.098)	(0.083)	(0.072)	(0.070)
P-value	0.440	0.944	0.525	0.527	0.241	0.987	0.165
Stepdown P-value	0.948	0.998	0.950	0.950	0.820	0.998	0.743
Comparison mean	0.981	-0.018	-0.079	-0.130	0.020	-0.105	16.562
N	1,654	1,528	1,528	941	1,528	1,524	1,522
P-val male=fem	0.851	0.259	0.576	0.678	0.019	0.610	0.363

The unit of observation for columns 1 is the child and the data comes from surveys with scholarship-lottery participants. The unit of observations for columns 2 to 5 is a caregiver-child pair and the data comes from surveys with caregivers. Standard errors clustered at the scholarship-lottery participant level. *** p<0.01, ** p<0.05, * p<0.1. All regressions control for baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant, and birth order of the child. Columns 2-5 also control for child age in months and child gender. Refer to Appendix B for components of the 'Preventive health behaviors index' (Table B2), 'Child investment index' (Table B5), 'Schooling index' (Table B6), and 'Child stimulation index' (Table B7). For the 'Schooling index', the sample is restricted to caregiver-child pairs with children over 36 months old. 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 B8). "Aspiration": shows the answer to the question "What is the highest level of education that you would like [child name] to complete?".

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 3 of Table 5, with breakdown shown in Table B5) nor on a schooling index (column 4 of Table 5, with breakdown shown in Table B6). However, among children of female scholarship recipients, the caregiver reports more interacting with the child in stimulating ways (column 5 of Table 5, with breakdown in Table B7).

Interestingly, we see this difference in reported behavior despite no effect on *knowledge* about the role of parental stimulation in children's cognitive development (columns 7 and 8 of Table 5). We also see no significant differences in caregiver's aspirations for their children—likely because they all believe "the sky is the limit," with an average expected 16 years of education.

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
Stepdown P-value	0.057	0.023	0.567	0.810	0.810	0.867
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
Stepdown P-value	0.207	0.147	0.147	0.018	0.043	0.147
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

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. "Treatment" means the child's parent won the scholarship lottery for Senior High School (SHS). 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, 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. *** p<0.01, ** p<0.05, * p<0.1

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).[27]

[27] We have non-trivial attrition in the LENA measurements, as shown in Table A2, but we find no imbalance

Summing up, we do not find clear evidence that female scholarship recipients 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 A10). 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. The negative LENA results for children of male scholarship recipients still hold with entropy balancing (Table A15).

6 Cost-effectiveness, Marginal Value of Public Funds and Benefit-Cost Ratio

Intergenerational impacts rarely factor into policy debates around subsidizing secondary education. Yet, our findings of substantial intergenerational impacts 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 marginal value of public funds (Hendren and Sprung-Keyser, 2020) 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).

in child characteristics (Table A4) nor in 2008 baseline characteristics or current household environment (Table A14) 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 A15.

6.1 Cost-effectiveness

We first compute the cost per child death averted and the cost per standard deviation improvement in child cognitive development. For this exercise, we consider only the scholarship costs and estimated administrative costs associated with the program if it were implemented at scale. We estimate a cost of $565 per scholarship recipient for four years of SHS. Since we are looking at impacts on *second-generation children*, 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. Given this, we also show in Panel C the cost-effectiveness when we impute a total fertility of 2.1 (the replacement rate). This is still much below the current total fertility rate of 3.4 for Ghana.

Combining these cost estimates with the child mortality reduction (Table 2) gives us a cost per under-3 death averted of around $23,000 for scholarships given to both genders, and of $17,500 for scholarships given to females only (column 1 Panel A of Table 7). Both of these estimates go down to around $15,000 when we take into consideration future fertility. 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 neonatal 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 intervention recommended by WHO 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.[28] Thus, our estimates imply that scholarship could be in the recommended set of interventions to reduce child mortality, even if they had no other impacts.

In Column 2 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 the intergenerational impacts of the scholarships on cognitive scores for seven year-olds (Table 3). Using this estimate yields a cost per SD increase in early childhood cognitive test scores of $2,872 if the program covers both genders and $1,208 if it covers females only (Table 7).

[28]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.

Table 7: Cost-Effectiveness, Marginal Value of Public Funds (MVPF), and Benefit-Cost Ratio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cost-effectiveness		**Marginal value of public funds**			**Benefit-cost ratio**		
			Value of a statistical life:					
	USD per death averted	USD per s.d of cognitive gain	Low	Medium	High	Low	Medium	High
Panel A: 2nd generation benefits								
Scholarships for all	23,338	2,872	0.9	3.1	4.9	0.7	2.5	4.1
	[14,935, 41,192]	[1,121, 7,652]	[0.6, 1.2]	[1.8, 4.2]	[2.8, 6.8]	[0.4, 0.9]	[1.4, 3.5]	[2.3, 5.5]
Scholarships for girls	17,595	1,208	1.1	3.8	6.1	0.9	3.2	5.2
	[10,404, 39,411]	[717, 2,751]	[0.6, 1.6]	[1.9, 5.6]	[2.9, 8.9]	[0.4, 1.3]	[1.5, 4.6]	[2.5, 7.4]
Panel B: 1st and 2nd generation benefits								
Scholarships for all			1.2	3.4	5.2	0.9	2.8	4.3
			[0.7, 1.5]	[2.0, 4.6]	[3.0, 7.1]	[0.5, 1.3]	[1.6, 3.8]	[2.4, 5.8]
Scholarships for girls			1.5	4.2	6.5	1.3	3.6	5.5
			[0.9, 2.1]	[2.1, 6.1]	[3.2, 9.4]	[0.7, 1.7]	[1.8, 5.1]	[2.7, 7.9]
Panel C: 1st and 2nd generation benefits with full fertility								
Scholarships for all	16,807	2,068	1.4	4.1	6.4	1.1	3.4	5.3
	[10,755, 29,663]	[808, 5,510]	[0.8, 1.8]	[2.4, 5.7]	[3.6, 8.8]	[0.6, 1.5]	[1.9, 4.7]	[3.0, 7.3]
Scholarships for girls	15,548	1,067	1.6	4.6	7.0	1.3	3.9	6.0
	[9,193, 34,824]	[633, 2,431]	[0.9, 2.2]	[2.3, 6.6]	[3.4, 10.2]	[0.7, 1.9]	[1.9, 5.5]	[2.9, 8.6]

Notes: 10% bootstrapped confidence intervals are shown in brackets. Panel A considers only the second-generation benefits from averting child deaths (Table 2). Panel B includes the benefits in Panel A, as well as the first-generation benefits. The number of births used for Panels A and B is equal to the sample births in the treatment group as of 2023. Panel C includes the same benefits as Panel B but assumes a total fertility rate of 2.1 children for both men and women. Columns 3 to 8: For the "Medium" columns, we use 100×GNI per capita in Ghana for the value of a statistical life (VSL). In the "Low" columns, we present a less optimistic scenario using 33×GNI per capita. In the "High" column, we present a more optimistic scenario using 160×GNI per capita. Refer to Appendix F for more details on these calculations.

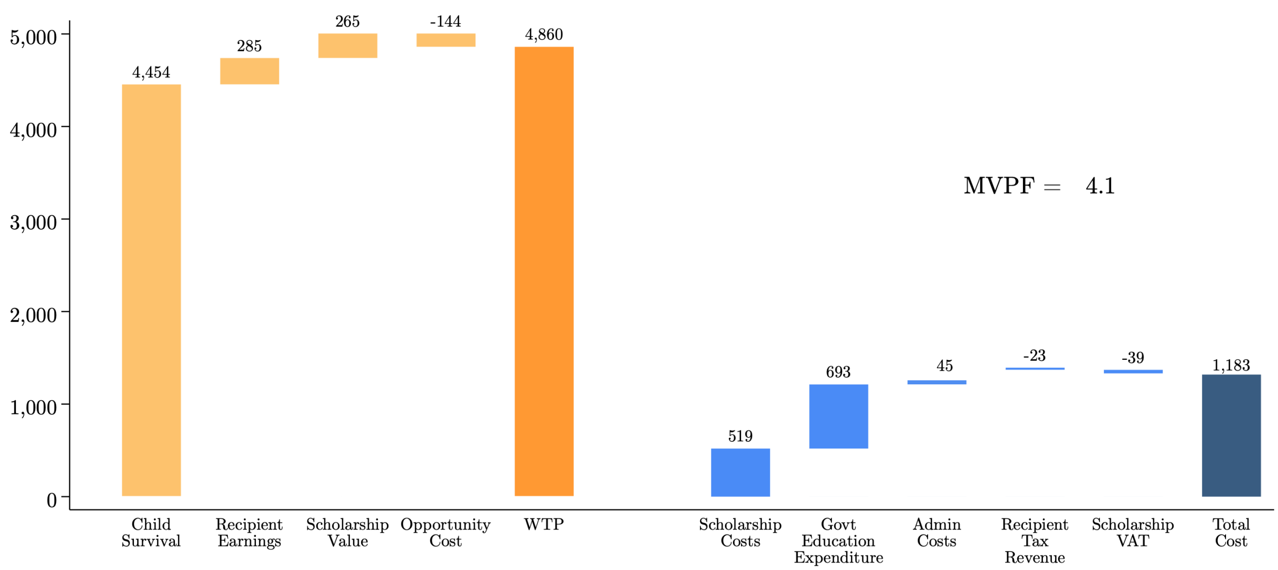
6.2 Marginal Value of Public Funds (MVPF) and Benefit-Cost Ratio

This subsection estimates the Marginal Value of Public Funds (MVPF) following the methodology proposed by Hendren and Sprung-Keyser (2020), as well as the more traditional Benefit-Cost Ratio (BCR). While the BCR considers the upfront government expenditure, the MVPF focuses on the long-run benefits per dollar of *net* government expenditure. Specifically, the MVPF is calculated as the ratio of two numbers: the benefits that a policy provides to its recipients divided by the policy's costs *net* of long-run government savings such as additional tax revenue raised from higher earners. The components of each of these two numbers are shown in Figure 3, with benefits in orange and costs in blue. We detail how each of these are computed/estimated in Appendix F.

MVPF: Benefits The numerator for the MVPF—the dark orange bar in Figure 3—adds the ways in which the policy improves the lives of its beneficiaries—Hendren and Sprung-Keyser

(2020) call this sum the individuals' "Willingness to Pay (WTP)".[29] We measure the benefits of the scholarship program from averted child deaths and greater earnings for the scholarship recipients. To be conservative, we ignore the value of early childhood cognitive gains.[30] We also 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).

Figure 3: Means-Tested Scholarships for All with Full Fertility: Breakdown of MVPF



Notes: Estimates correspond to column 4, row 1, panel C of Table 7. The figure shows in light orange the components of the *numerator* (dark orange) and in blue the components of the *denominator* (dark blue) in the marginal value of public funds (MVPF) calculation. The costs presented assume a scaled implementation of the program. Administrative costs as incurred by the research team were 181 USD per beneficiary, leading to a total net cost of 1319 USD and an MVPF of 3.7. See details in Appendix F.

We calculate the total benefits from averted child deaths using estimates of the value of statistical life (VSL) for our sample. We follow Robinson et al. (2019)'s recommendation for conducting VSL sensitivity analyses. 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 VSL' column in Table 7), $76,862 (refer to 'Low VSL' column), and $376,000

[29]This is what individuals would be willing to pay for the program if they were maximizing their lineage well-being, absent any market or information frictions.

[30]If one wanted to quantify those, one would need to project the effect of these gains on adult earnings and taking the present value of this income stream. This is challenging because there are no estimates of the relationship between early childhood cognitive development and adult earnings in our context.

(refer to 'High VSL' column), respectively. Note that all these VSLs (even the "high" one) are very low compared to the numbers used in the United States.[31]

Since Duflo et al. (2025) find no impact on earnings until 2020, we model the first-generation labor market benefits as a stream of increased income beginning at age 30 and continuing for the remaining 34 years of working life. Following Duflo et al. (2025), we attribute the observed earnings gains to a 6.7 percentage point increase in public sector employment for women. Thus, we ignore any potential within-sector wage increases. To estimate the impact, we calculate average wages in the public and non-public sectors by gender using sample data, and project earnings over time using ILO lifetime earnings profiles for Ghana. We then calculate the annual earnings increase for public sector workers, and multiply the net present value of this stream by the program's estimated ITT effect on public sector employment.

MVPF: Costs The denominator for the MVPF (dark blue bar labeled "Total Cost") departs from the cost value used in the cost-effectiveness section above in that it (a) is net of tax revenues gained in the long run, and (b) it includes *all* associated costs incurred by the government, not only the scholarship costs. In particular, we include (i) the cost of administering a means-tested scholarship program, including the cost of identifying eligible youths;[32] (ii) the per-student government education expenditure in secondary school (this includes, for example, teacher salaries and lab equipment), and (iii) the per-student government education expenditure in tertiary programs (since some of the secondary school scholarship recipients go on to tertiary).

MVPF The MVPF estimates are shown in columns 3-5 of Table 7. In Panel A, we consider the second generation survival benefits, considering only children born within the study period. In Panel B, we add the first-generation labor market benefits estimated in Duflo et al. (2025). In Panel C, we consider the second generation survival benefits assuming 2.1 children per scholarship recipient (a conservative estimate since the average total fertility rate in Ghana was estimated at 3.5 for 2022) and the first-generation labor market benefits.

Our calculations indicate that secondary school scholarships produce substantial benefits,

[31] For example, the Environmental Protection Agency (EPA)'s website recommends using a VSL of USD7.4 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.)

[32] To select scholarship eligible students, Duflo et al. (2025) conducted school and home visits after the start of the school year to identify 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. At scale, we would expect a government to identify eligible youths by exploiting administrative records and seeking inputs from school principals. We consider two costs scenarios: the administration costs incurred during the study ("pilot") and the administration costs that we estimate would be incurred by the government at "scale").

with an at-scale MVPF of 4.1 (meaning that each dollar spent on the policy generates 4.1 dollars in benefits) in the medium VSL scenario. The MVPF is higher (4.6) when scholarships are targeted to female students, and higher still if we consider the higher VSL value (7.0).

Importantly, the benefits of secondary education for the next generation outstrip the benefits to the first-generation. As one can see by comparing Panels A and B of Table 7, the benefits to the second-generation account for almost all of the numerator in the MVPF. This is important to note, since the inter-generational impacts are likely to be much less sensitive to possible general equilibrium impacts than the first-generation impacts. Duflo et al. (2025) 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.

6.3 Benefit-Cost Ratio

We provide estimates of the benefit-cost ratio as more typically estimated in the public health literature in columns 6-8 of Table 7. The details are provided in Appendix F. Under the medium VSL valuation, and considering total expected fertility, the benefits of scholarships for all are 3.4 times their costs.

7 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.[33] 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.

[33]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.

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.

Impacts on the children of male scholarship recipients are harder to estimate due to data censoring: even after 15 years, only about half of the men in the sample report having had a child, and there appears to be heterogeneous impacts in household formation, limiting our ability to make strong causal claims at this time. What appears clear is that there are no systematic, large positive impacts such as those observed for children of female scholarship participants within that time frame. Duflo et al. (2025) 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 woman induced to attend secondary school by the scholarship would have higher returns to secondary school than the marginal man. Future work is needed to ascertain whether the non-labor market returns for the marginal men were also lower than for the marginal women.

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.

Carneiro, Pedro, Costas Meghir, and Matthias Parey, "Maternal education, home environments, and the development of children and adolescents," *Journal of the European Economic Association*, 2013, *11* (suppl_1), 123–160.

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*, 2025.

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," *American Economic Journal: Economic Policy*, 2025, *17* (2), 388–417.

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.

Fernald, Lia CH, Elizabeth Prado, Patricia Kariger, and Abbie Raikes, "A toolkit for measuring early childhood development in low and middle-income countries," 2017.

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

Finkelstein, Amy and Nathaniel Hendren, "Welfare analysis meets causal inference," *Journal of Economic Perspectives*, 2020, *34* (4), 146–167.

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.

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.

Hendren, Nathaniel and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 2020, *135*, 1209–1318.

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.

Källberg, David and Ingeborg Waernbaum, "Large Sample Properties of Entropy Balancing Estimators of Average Causal Effects," *Econometrics and Statistics*, 2023.

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.

Walker, Michael W, Alice H Huang, Suleiman Asman, Sarah J Baird, Lia Fernald, Joan Hamory Hicks, Fernando Hoces de la Guardia, Satoshi Koiso, Michael Kremer, Matthew N Krupoff et al., "Intergenerational child mortality impacts of deworming: Experimental evidence from two decades of the Kenya Life Panel Survey," Technical Report, National Bureau of Economic Research 2023.

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.

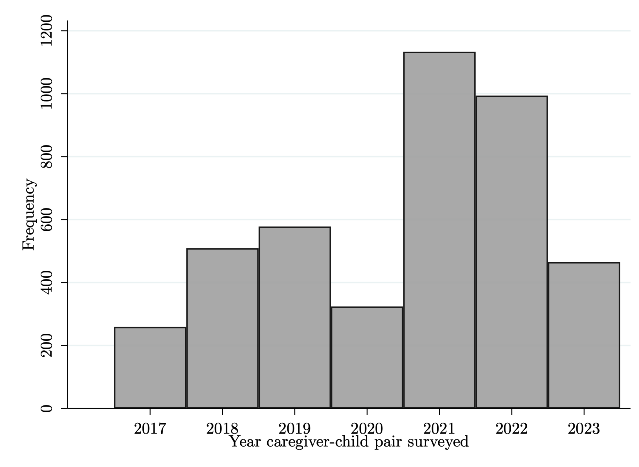
Xu, Dongxin, Umit Yapanel, and Sharmi Gray, "Reliability of the LENA Language Environment Analysis System in young children's natural home environment," *Boulder, CO: Lena Foundation*, 2009, pp. 1–16.

Online Appendix

Table of Contents

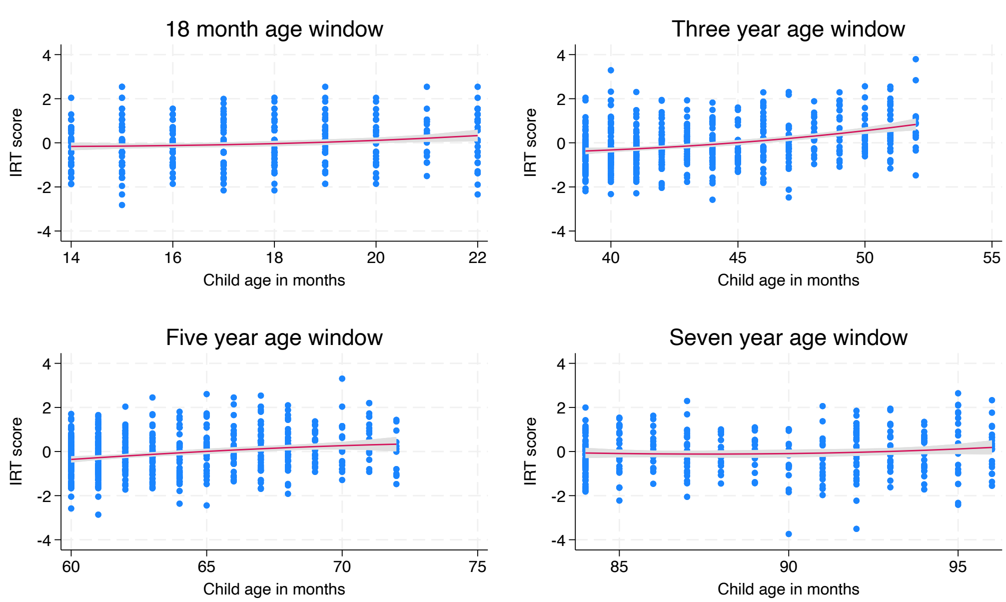
A Appendix Figures and Tables

Figure A1: 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.

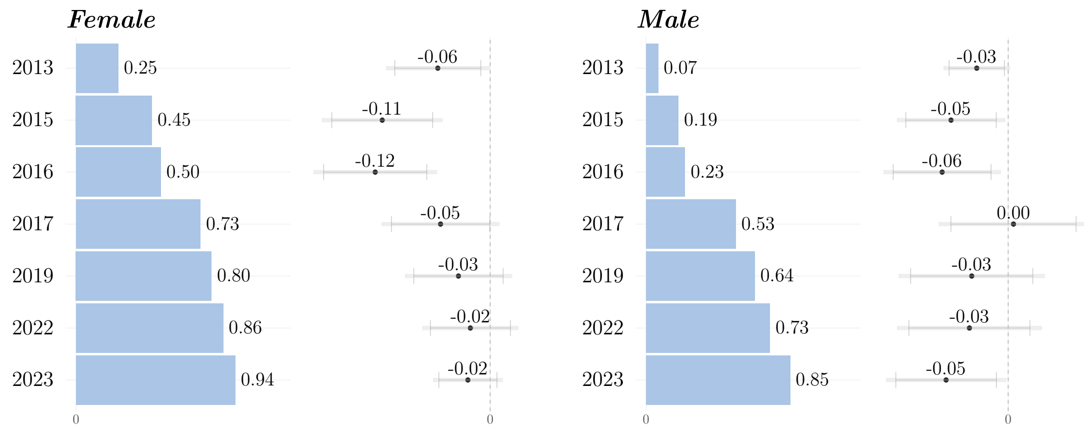
Figure A2: Cognitive Development Scores by Age in Months



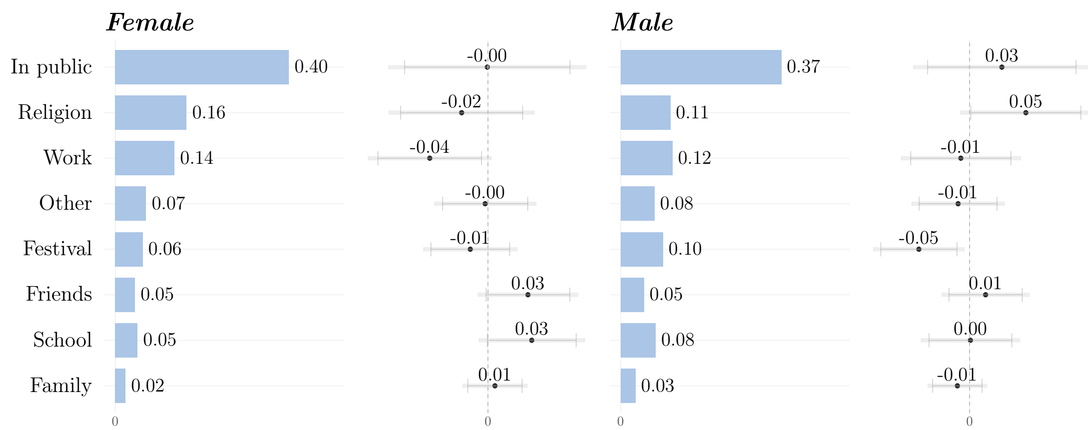
Notes: Plots only include the control group's children. Data are from 2017-2022 in-person measures administered by the surveyors with the target children at the following target ages: 14-22 months old ("18 month"), 39-52 months old ("Three"), 60-72 months old ("Five"), 84-96 months old ("Seven"). The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

Figure A3: Marriage Market Outcomes for Scholarship-Lottery Participants

(a) Share ever living with a partner



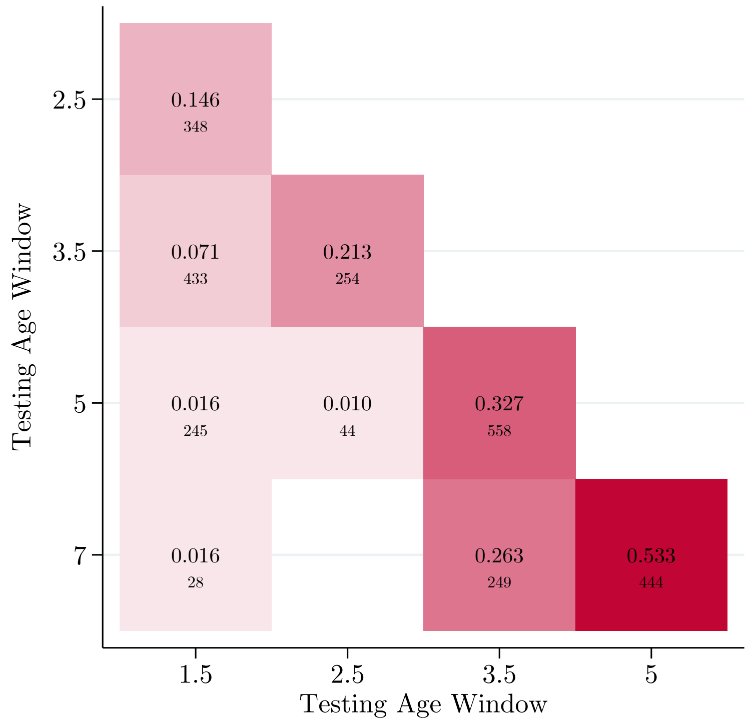
(b) How/where did you meet your most recent partner?



(c) Age Gap between Male and Female Partner

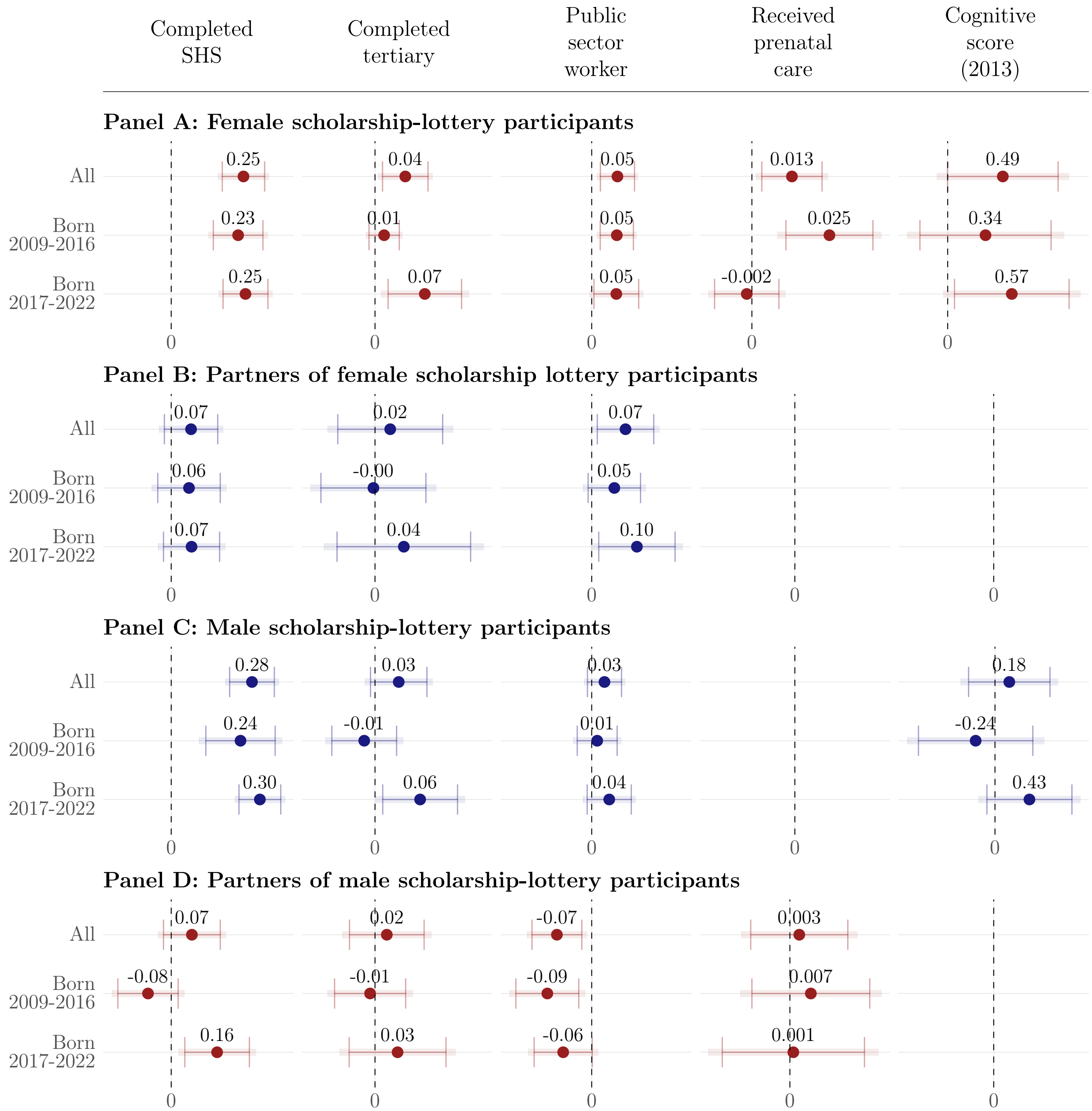
Notes: Observations are scholarship-lottery participants. Bars show comparison group means, and the point estimates and confidence intervals for the treatment effects. Subfigure (b) includes participants ever living with a partner by 2022.

Figure A4: 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.

Figure A5: Scholarship lottery participant and partner characteristics: Survival-to-1 sample



Notes: An observation is a second-generation child for whom survival to age 1 is available. The graphs show characteristics of the child's "scholarship eligible" parent (Panel A if female, Panel C if male) and of the partner of the child's "scholarship-eligible" parent (Panels B and D). Within each panel, the two row shows the results for the pooled sample, and the next two rows split the sample at the median birth year (2017).

Table A1: First Generation Impact of Scholarship—scholarship-lottery participants with at least one child

	(1) Cognitive test score (2013)	(2) Completed SHS (2019)	(3) Completed tertiary (2023)	(4) Earnings last 6 months (GHX, 2019)	(5) Earnings last 6 months (GHX, 2023)	(6) Self-employed (2019)	(7) In public sector (2019)	(8) In public sector (2023)
Panel A: Female scholarship-lottery participants								
Treatment	0.182**	0.242***	0.100***	62.080	558.113**	0.003	0.039*	0.060***
	(0.080)	(0.037)	(0.029)	(119.138)	(255.201)	(0.036)	(0.020)	(0.023)
P-value	0.023	0.000	0.001	0.602	0.029	0.927	0.053	0.010
Stepdown P-value	0.125	0.000	0.005	0.855	0.125	0.927	0.152	0.072
Comparison mean	-0.260	0.331	0.088	901.753	1768.683	0.317	0.048	0.054
N	792	789	688	781	685	789	789	687
Panel B: Male scholarship-lottery participants								
Treatment	0.125	0.314***	0.020	214.377	122.011	-0.030	0.019	0.026
	(0.084)	(0.042)	(0.030)	(197.551)	(406.727)	(0.039)	(0.023)	(0.028)
P-value	0.137	0.000	0.505	0.278	0.764	0.437	0.417	0.360
Stepdown P-value	0.618	0.000	0.882	0.831	0.882	0.882	0.882	0.882
Comparison mean	0.040	0.409	0.101	2106.234	4542.702	0.262	0.063	0.085
N	549	557	491	542	480	557	557	489
P-val male=fem	0.572	0.145	0.069	0.450	0.346	0.613	0.445	0.277

Notes: Observations are scholarship-lottery participants who ever had a child by 2023, and are in the child survival sample. Data Sources: surveys conducted in 2013, 2019, and 2023. 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; the third and fourth row show the analytical p-value and the step-down p-value correcting for multiple hypothesis testing; control group means are in each panel's fifth row; sample size for the estimation is in each panel's sixth 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						
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
5 yrs old sample	0.90	0.90	836	0.90	359	0.958
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						
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
5 yrs old sample	0.97	0.97	348	0.97	179	0.762
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 only those born at least 12 months prior to 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	-0.172	0.004	-0.022	-0.408**	-0.005	0.020
	(0.133)	(0.007)	(0.043)	(0.181)	(0.042)	(0.031)
P-value	0.197	0.535	0.603	0.024	0.899	0.512
Comparison mean	17.470	0.617	0.468	5.565	0.503	0.151
N	819	760	814	819	815	815
Male Scholarship-lottery participants:						
Treatment	-0.238	0.007	0.097*	-0.248	0.037	-0.094***
	(0.185)	(0.008)	(0.050)	(0.242)	(0.050)	(0.032)
P-value	0.199	0.366	0.055	0.305	0.455	0.004
Comparison mean	17.600	0.624	0.381	5.803	0.503	0.194
N	569	535	567	569	564	564
P-val male=fem	0.768	0.822	0.073	0.578	0.525	0.012
Panel B: 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
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)	69.74	40.00	69.75	39.82	530	69.74	40.37	299	0.998
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: Main results Pooling Female and Male Scholarship Lottery Participants

	(1)	(2)	(3)	(4)	(5)	(6)
	Survival			Cognitive ability index score		
Age:	1 years	3 years	5 years	3 year	5 years	7 years
Treatment	0.016**	0.016**	0.016	0.021	0.101	0.130
	(0.006)	(0.008)	(0.011)	(0.058)	(0.071)	(0.098)
P-value	0.012	0.036	0.131	0.723	0.158	0.182
Stepdown P-value	0.022	0.044	0.132	0.750	0.398	0.398
Comparison mean	0.967	0.962	0.954	-0.004	-0.001	-0.001
N	2,716	2,160	1,599	970	964	535

*** p<0.01, ** p<0.05, * p<0.1
Notes: See Table 2.

Table A6: Survival Sample: Robustness to Choice of Controls and Mother's age at birth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Survival up to one year				Mother's age at birth		
					All	Born 2009-2016	Born 2017-2022
Panel A: Children of female scholarship-lottery participants							
Treatment	0.017**	0.017**	0.018**	0.019**	0.345	0.064	0.063
	(0.008)	(0.008)	(0.008)	(0.008)	(0.238)	(0.207)	(0.168)
P-value	0.033	0.039	0.031	0.023	0.147	0.756	0.706
Comparison mean	0.965	0.965	0.965	0.965	24.797	22.112	28.709
N	1,707	1,707	1,707	1,707	1,825	1,065	760
Panel B: Children of male scholarship-lottery participants							
Treatment	0.014	0.013	0.013	0.013	-0.460	-0.559	-0.516
	(0.010)	(0.010)	(0.010)	(0.010)	(0.376)	(0.490)	(0.338)
P-value	0.162	0.164	0.174	0.161	0.222	0.255	0.127
Comparison mean	0.970	0.970	0.970	0.970	23.022	20.285	25.070
N	1,016	1,016	1,016	1,016	1,033	436	597
P-val male=fem	0.713	0.751	0.730	0.700	0.113	0.314	0.136
Birth order fixed effects	Yes	No	Yes	Yes	Yes	Yes	Yes
Mother's age at birth control	No	No	Yes	No	No	No	No
Linear birth year control	Yes	Yes	Yes	No	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)-(7) control for the JHS finishing exam score and baseline region fixed effects of the scholarship-lottery participant, as well as the birth order of the child.

Table A7: 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
Stepdown P-value	0.945	0.951	0.951	0.033	0.136
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
Stepdown P-value	0.663	0.284	0.920	0.284	0.920
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 A8: 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)
Female child	0.153	-0.095	0.167**	-0.084	0.247**
	(0.108)	(0.153)	(0.083)	(0.085)	(0.124)
Treatment x Female child	-0.282	0.077	-0.011	0.219	-0.159
	(0.187)	(0.263)	(0.139)	(0.162)	(0.222)
P-value	0.553	0.749	0.750	0.255	0.030
Stepdown P-value	0.911	0.929	0.929	0.676	0.144
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)
Female child	0.216	-0.099	0.124	0.043	-0.048
	(0.132)	(0.165)	(0.116)	(0.144)	(0.206)
Treatment x Female child	-0.192	-0.017	-0.272	0.318	0.612*
	(0.215)	(0.307)	(0.180)	(0.212)	(0.368)
P-value	0.099	0.395	0.265	0.009	0.070
Stepdown P-value	0.285	0.470	0.470	0.060	0.285
Comparison mean	-0.009	-0.034	0.041	-0.043	-0.114
N	342	207	342	298	174
P-val male=fem	0.322	0.259	0.918	0.003	0.113

*** p<0.01, ** p<0.05, * p<0.1
See Table 3 notes.

Table A9: 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
Stepdown P-value	0.962	0.962	0.889	0.962
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
Stepdown P-value	0.380	0.426	0.380	0.716
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 A10: Household Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Caregiver is Mother	Caregiver is Father	Caregiver is Grandmother	Lives with Mother	Lives with Father	Lives with both parents	Num. other children in household	Number of siblings	Number of adults in household
Panel A: Children of female scholarship-lottery participants									
Treatment	-0.003	-0.001	0.006	-0.009	0.004	0.007	0.059	-0.104	0.039
	(0.017)	(0.006)	(0.014)	(0.014)	(0.034)	(0.035)	(0.117)	(0.134)	(0.088)
P-value	0.841	0.891	0.691	0.515	0.906	0.851	0.615	0.437	0.657
Stepdown P-value	0.996	0.996	0.993	0.984	0.996	0.996	0.993	0.968	0.993
Comparison mean	0.906	0.014	0.061	0.933	0.646	0.632	1.642	1.926	2.380
N	3,087	3,087	3,087	3,042	2,695	2,669	2,757	2,996	2,733
Panel B: Children of male scholarship-lottery participants									
Treatment	0.036	-0.021	-0.022	0.024	-0.087**	-0.092**	0.160	0.174	0.060
	(0.024)	(0.018)	(0.015)	(0.019)	(0.039)	(0.043)	(0.133)	(0.142)	(0.098)
P-value	0.136	0.238	0.138	0.215	0.024	0.032	0.227	0.223	0.544
Stepdown P-value	0.574	0.699	0.574	0.699	0.179	0.223	0.699	0.699	0.699
Comparison mean	0.740	0.158	0.077	0.888	0.711	0.656	1.323	1.416	2.369
N	1,767	1,767	1,767	1,603	1,647	1,491	1,536	1,686	1,523
P-val male=fem	0.138	0.280	0.133	0.114	0.115	0.120	0.604	0.170	0.965

*** 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 row reports the p-values; the fourth row reports the family-wise Romano-Wolf step-down p-values; comparison group means are in the fifth row; the sixth row reports the number 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 A11: Main results: Robustness to including partner's education

	(1)	(2)	(3)	(4)	(5)
		Survival		Cognitive ability index score	
Age:	1 years	3 years	5 years	5 years	7 years
Panel A: Children of Female scholarship-lottery participants					
Treatment	0.018**	0.022**	0.026*	0.218***	0.254**
	(0.008)	(0.010)	(0.013)	(0.083)	(0.122)
Most recent partner completed SHS	-0.005	-0.005	-0.004		
	(0.011)	(0.013)	(0.015)		
Most recent partner completed tert.	-0.003	-0.009	-0.024		
	(0.013)	(0.017)	(0.023)		
Partner of caregiver completed SHS				0.258***	0.070
				(0.087)	(0.134)
Partner of caregiver completed tert.				0.034	0.039
				(0.127)	(0.193)
P-value	0.027	0.032	0.052	0.009	0.038
Comparison mean	0.965	0.958	0.947	0.023	0.064
N	1,707	1,395	1,084	657	353
Panel B: Children of Male scholarship-lottery participants					
Treatment	0.014	0.007	-0.000	-0.205*	-0.042
	(0.010)	(0.012)	(0.017)	(0.124)	(0.193)
Most recent partner completed SHS	0.000	0.006	-0.003		
	(0.011)	(0.013)	(0.021)		
Caregiver completed SHS				0.366**	0.621**
				(0.165)	(0.293)
P-value	0.151	0.568	0.990	0.098	0.829
Comparison mean	0.970	0.971	0.968	-0.043	-0.114
N	1,016	772	520	297	172

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

Notes: Columns (1), (2), and (3): 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 (4) and (5): an observation is a child of a participant in the lottery for secondary school scholarships at a given age window. 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 ares 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 A12: Child Mortality, restricted to firstborns

	(1)	(2)	(3)
	Survival		Mother's age at birth
	Up to 1 year	Up to 3 years	
Panel A: Children of female scholarship-lottery participants			
Treatment	0.010	0.013	0.549*
	(0.015)	(0.016)	(0.307)
P-value	0.506	0.415	0.074
Stepdown P-value	0.510	0.478	0.160
Comparison mean	0.958	0.952	22.272
N	691	635	798
Panel B: Children of male scholarship-lottery participants			
Treatment	0.006	0.002	0.128
	(0.014)	(0.016)	(0.376)
P-value	0.639	0.900	0.734
Stepdown P-value	0.913	0.951	0.951
Comparison mean	0.974	0.972	21.501
N	522	440	539
P-val male=fem	0.796	0.648	0.409

*** 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 row reports the p-values; the fourth row reports the family-wise Romano-Wolf step-down p-values; comparison group means are in the fifth row; the sixth row reports the number 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 A13: 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
Stepdown P-value	0.193	0.429	0.925	0.764
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
Stepdown P-value	0.913	0.522	0.843	0.735
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 B3 and Table B4. Refer to Table 4 for other details on specifications and outcomes.

Table A14: 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 A15: 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	0.332***	0.069***	0.009	0.378	0.381	-0.003
	(0.120)	(0.023)	(0.007)	(0.709)	(0.599)	(0.279)
P-value	0.005	0.002	0.234	0.594	0.525	0.991
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	-0.210	-0.067**	-0.019**	-2.999***	-2.390***	-0.609*
	(0.148)	(0.028)	(0.008)	(0.827)	(0.707)	(0.312)
P-value	0.156	0.016	0.017	0.000	0.001	0.051
Comparison mean	2.210	0.380	0.171	14.260	10.677	3.584
N	391	391	391	391	391	391

*** 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) Mud walls (reversed in index)	(4) Metal sheet roof	(5) Phone credit value (GHS, last 7 days)
Panel A: Children of female scholarship-lottery participants					
Treatment	0.014	-1.135	-0.027	0.010	-0.420
	(0.014)	(3.254)	(0.025)	(0.012)	(0.925)
P-value	0.319	0.727	0.284	0.420	0.650
Stepdown P-value	0.801	0.866	0.801	0.801	0.866
Comparison mean	0.403	74.497	0.151	0.960	11.706
N	2,729	2,734	2,741	2,741	2,728
Panel B: Children of male scholarship-lottery participants					
Treatment	0.021	-2.539	-0.041	-0.007	-0.597
	(0.019)	(4.365)	(0.037)	(0.013)	(1.023)
P-value	0.257	0.561	0.267	0.616	0.560
Stepdown P-value	0.780	0.922	0.780	0.922	0.922
Comparison mean	0.409	86.255	0.241	0.966	13.597
N	1,518	1,523	1,525	1,525	1,523
P-val male=fem	0.758	0.743	0.810	0.368	0.671

*** 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-lottery participant level are in the second row in parentheses; the third row reports the p-values; the fourth row reports the family-wise Romano-Wolf step-down p-values; comparison group means are in the fifth row; the sixth row reports the number 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.

Table B2: Preventive Health Behaviors—Index Components

	(1) Took child for check-up in past 12 mo	(2) Child sleeps under mosquito net	(3) Treats child's drinking water	(4) Shows card and has all vaccines	(5) Improved Toilet	(6) HH has priv. toilet
Panel A: Children of female scholarship-lottery participants						
Treatment	0.011	0.057*	-0.002	0.013	-0.004	0.076**
	(0.020)	(0.034)	(0.012)	(0.028)	(0.033)	(0.034)
P-value	0.598	0.093	0.851	0.630	0.900	0.027
Stepdown P-value	0.970	0.412	0.979	0.970	0.979	0.175
Comparison mean	0.368	0.610	0.047	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	-0.003	-0.024	-0.004	-0.012
	(0.025)	(0.042)	(0.018)	(0.036)	(0.039)	(0.039)
P-value	0.296	0.115	0.865	0.497	0.912	0.753
Stepdown P-value	0.829	0.547	0.981	0.930	0.981	0.981
Comparison mean	0.400	0.588	0.052	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	0.953	0.327	0.859	0.070

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

Notes: Refer to Table B1.

Table B3: 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
Stepdown P-value	0.882	0.356	0.882	0.641	0.882	0.882
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
Stepdown P-value	0.893	0.980	0.980	0.980	0.980	0.388
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 B4: 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
Stepdown P-value	0.830	0.901	0.949	0.933
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
Stepdown P-value	0.052	0.236	0.236	0.236
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 B5: 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
Stepdown P-value	0.564	0.564	0.847	0.770
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
Stepdown P-value	0.888	0.888	0.888	0.824
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 B6: 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
Stepdown P-value	0.904	0.904	0.904	0.895
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
Stepdown P-value	0.715	0.958	0.846	0.715
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.

Table B7: 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
Stepdown P-value	0.066	0.454	0.243	0.121	0.087
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
Stepdown P-value	0.756	0.802	0.694	0.065	0.802
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 B8: 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
Stepdown P-value	0.461	0.495	0.669	0.495
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
Stepdown P-value	0.700	0.765	0.700	0.700
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.

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
Stepdown P-value	0.664	0.869	0.847	0.025	0.162
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
Stepdown P-value	0.624	0.512	0.956	0.278	0.850
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
Stepdown P-value	0.835	0.146	0.061
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
Stepdown P-value	0.754	0.005	0.754
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
Stepdown P-value	0.959	0.495	0.188	0.106
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
Stepdown P-value	0.683	0.773	0.840	0.965
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
Stepdown P-value	0.700	0.517	0.104	0.223
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
Stepdown P-value	0.718	0.718	0.327	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
Stepdown P-value	0.862	0.967	0.917	0.017	0.437
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
Stepdown P-value	0.659	0.723	0.659	0.914	0.459
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
Stepdown P-value	0.991	0.991	0.991	0.991
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
Stepdown P-value	0.747	0.912	0.664	0.664
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.

1

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.

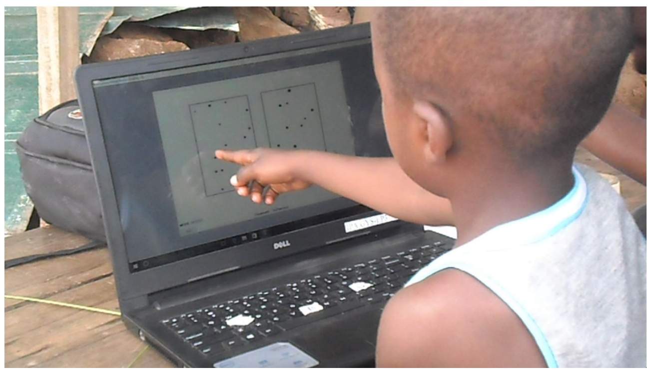
Figure D.1: 3-year old playing Vocabulary game

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:[34]

- **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).

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

- **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

F Details of the Cost-Effectiveness, Cost-Benefit, and MVPF Analyses

This section describes in-detail how we calculate the cost-effectiveness, benefit-cost ratios, and MVPF estimates presented in Table 7.

F.1 Benefits

We consider the benefits from the scholarship for the recipients in terms of averted child deaths and increases in scholarship recipient wages. The numerator in the MVPF calculations includes the value of the scholarships themselves. The positive fiscal externalities for the government include increased tax revenue and VAT from the scholarships.

Child survival

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.

We estimate the total survival-related benefits generated by the program by multiplying the value of a statistical life (VSL), adjusted for age weighting and the estimated program impact on child survival, by the average number of children per participant in the treated group.

We define three cohorts of children:

- **Cohort 1 (born by 2017)**: These children's survival benefits are discounted by 10 years.
- **Cohort 2 (born 2017–2023)**: Discounted by 15 years.
- **Cohort 3 (post-2023 to reach replacement fertility)**: Only included in the full fertility scenario. Discounted by 19 years.

Total survival benefits are indexed by:

- $s \in \{\text{low}, \text{medium}, \text{high}\}$: VSL scenario,
- p: program group (e.g., girls-only, all),
- $f \in \{\text{not-full}, \text{full}\}$: fertility scenario.

The total survival benefits expression is:

$$\text{Total Survival Benefits}_{p,f,s} = \text{VSL}_s \cdot \text{Age weight} \cdot \beta_{\text{survival}}(p) \cdot \left(\frac{N^{(1)}_{\text{child},p}}{(1+0.05)^{10}} + \frac{N^{(2)}_{\text{child},p}}{(1+0.05)^{15}} + \mathbf{1}_{\{f=\text{full}\}} \cdot \frac{N^{(3)}_{\text{child},p}}{(1+0.05)^{19}} \right)$$

where:

- $\beta_{\text{survival}}(p)$ is the estimated program impact on child survival under program p,
- $N^{(1)}_{\text{kids},p}$: Average number of children ever born in treatment group by 2017,
- $N^{(2)}_{\text{kids},p}$: Average number of children ever born in treatment group between 2017 and 2023
- $N^{(3)}_{\text{kids},p}$: Additional children needed to reach replacement fertility (2.1), only included for full fertility scenario ($f = \text{full}$).

We multiply the VSL by an age weight of 1.33 to reflect the fact that more years of life are lost due to under-3 deaths relative to deaths of the average Ghanaian.[35]

$\beta_{\text{survival}}(p)$ differs by the program type. Given our assumptions, $\beta_{\text{survival}}(p)$ is equal to the effect on the children of female scholarship recipients (1.7 pp increases; 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 increase) when we evaluate the program open to both genders.

Earnings and Income Tax

Following Duflo et al. (2025), we attribute the observed earnings gains among female scholarship recipients entirely to increased employment in the public sector: a 6.7 percentage point increase for women, and 3.9 percentage points when including both female and male recipients. Among women, this rise in public sector employment is primarily due to increases in teaching (4 percentage points) and health sector employment (2 percentage points).

To estimate the scholarship's impact on lifetime earnings, we model earnings separately for public sector and non-public sector workers, disaggregated by gender. For public sector workers, we assume a starting monthly wage equal to the average in our sample among public sector workers—USD 91 for women and USD 111 for men (2020 USD, nominal). For non-public sector workers, the starting wage is set to the monthly average earnings in the control group (USD 25 for women, USD 58 for men).

To project earnings over the recipients' working lives, we apply age- and sector-specific earnings growth factors from ILO data on the Ghanaian labor market. We then calculate the annual difference in earnings (net of income tax in the public sector), discount these differences back to the start of the program, and sum them over the working life. Finally, we multiply this sum by the estimated treatment effect on public sector employment to quantify the program's impact on lifetime after-tax earnings:

$$\text{Lifetime Earnings Impact}_p = \beta_{\text{public sector}}(p) \cdot \sum_{t=30}^{60} \frac{E^{\text{pub}}_t - E^{\text{non}}_t - T^{\text{pub}}_t}{(1+0.05)^{t-17}}$$

where $\beta_{\text{public sector}}(p)$ is the estimated treatment effect on public sector employment as a function of program type p (e.g., 0.067 for women only), E^{pub}_t and E^{non}_t are projected earnings in year t, T^{pub}_t is

[35] 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).

the income tax owed by public sector workers based on Ghana's income tax scheme, and 0.05 is the discount rate.

We assume full tax compliance in the public sector and zero compliance in the non-public sector. Accordingly, we calculate the program-induced increase in income tax revenue as:

$$\text{Lifetime Income Tax Revenue Impact}_p = \beta_{\text{public sector}}(p) \cdot \sum_{t=30}^{64} \frac{T_t^{\text{pub}}}{(1+0.05)^{t-17}}$$

This expression captures the fiscal return to the government under the assumption that income taxes are only collected from public sector workers.

Scholarship value, or Willingness to Pay (WTP) for Scholarships

The numerator of the Marginal Value of Public Funds includes recipients' willingness to pay (WTP) for secondary school scholarships. Following the methodology of Finkelstein and Hendren (2020), we calculate this as a weighted average of WTP among both marginal and inframarginal recipients. For inframarginal recipients—those who would have completed senior high school even in the absence of the scholarship—the program functions effectively as a cash transfer, and their WTP equals the full value of the scholarship. For marginal recipients—those whose schooling decisions were influenced by the program—we do not observe whether it was the first or last dollar of the subsidy that induced SHS completion. That is, even a scholarship worth less than full tuition might have been sufficient to alter their behavior. In the absence of precise data on this threshold, we adopt a midpoint assumption: we treat half of the scholarship value as the cost of inducing enrollment (the behavioral margin), and the other half as a cash-equivalent transfer.

The formula is given by:

$$\text{Program WTP}_p = \text{Share Inframarginal}_p \cdot \text{Scholarship Cost}_p + \frac{1}{2} \cdot \text{Share Marginal}_p \cdot \text{Scholarship Cost}_p$$

where:

- $\text{Share Inframarginal}_p$ is the share of control group participants that completed SHS under program p,
- Share Marginal_p is the treatment effect on SHS completion under program p,
- $\text{Scholarship cost}_p$ is the full monetary value of the scholarship.

VAT and Fiscal Externalities

The scholarship program affects government revenues through Ghana's value-added tax (VAT) system. We distinguish between two channels:

1. A direct fiscal offset: when disbursing scholarship funds, a portion of that spending returns immediately to the government through VAT, since scholarships are used to purchase VAT-liable goods and services such as tuition, boarding, and school materials;
2. A long-run fiscal externality: recipients who experience higher lifetime earnings also consume more over their lifetimes, generating additional VAT revenue.

Let:

- $\tau_{\text{VAT}} = 0.15$: the statutory VAT rate (15%),
- θ_{VAT}: the VAT compliance rate, (50%)
- Scholarship Cost_p: present value of the scholarship cost per recipient in program p,
- Lifetime Earnings Impact_p: increase in lifetime after-tax earnings caused by program p, as defined previously.

VAT offset from scholarship spending.
Because scholarships are spent on VAT-liable goods and services, part of the expenditure returns immediately to the government. This direct fiscal offset is given by:

$$\text{VAT Offset}_p = \tau_{\text{VAT}} \cdot \theta_{\text{VAT}} \cdot \text{Scholarship Cost}_p$$

VAT revenue from lifetime consumption.
Recipients with higher lifetime earnings also generate more consumption, a share of which is subject to VAT. The resulting long-run fiscal externality is:

$$\text{VAT Revenue from Lifetime Consumption}_p = \tau_{\text{VAT}} \cdot \theta_{\text{VAT}} \cdot \text{Lifetime Earnings Impact}_p$$

These expressions reflect two ways in which the VAT system reduces the net fiscal cost of the program: an immediate return of a portion of scholarship spending, and a long-run expansion of the VAT base through increased consumption.

F.2 Costs

Direct program costs

The most important cost of the program is the value of the senior secondary school scholarships. We calculate the total present value of scholarship costs by summing disbursements across the years 2009 to 2013 and discounting each to the program start year (2008). All values are assumed to be in constant 2020 USD:

$$\text{Scholarship Cost}_p = \frac{1}{N^{(p)}_{\text{treated}}} \sum_{y=2009}^{2013} \frac{C_{p,y}}{(1+0.05)^{t_y}}$$

Where:

- Scholarship Cost_p: Present value of scholarship cost per recipient in program p,
- $N^{(p)}_{\text{treated}}$: Number of scholarship recipients under program p,
- $C_{p,y}$: Total value of scholarships disbursed in year y under program p, in 2020 USD,
- $t_y = y - 2008$: Number of years from disbursement year y to base year 2008.

Indirect program costs

We divide indirect program costs into two components: *targeting costs* (from the listing survey used to identify eligible students) and *administrative costs* (staff time used to implement the program).

These are calculated under two scenarios: (i) as incurred during the pilot and (ii) under a scaled-up implementation.

Notation:

- Number of eligible students: $N_{\text{eligible}} = 2064$
- Share included in 2016 survey: $p_{\text{survey}} = 0.98$
- Listing cost: $C_{\text{total}}^{\text{listing}}$
- Let $N_{\text{treated}}^{(p)}$ be the number of scholarship recipients under program type p.
- Admin staff monthly salary: W

Targeting (Listing Survey) Costs: During the study, a full listing survey was conducted. Since only a subset of surveyed students received scholarships, we scale the listing cost by the proportion treated:

$$C_{\text{pilot}}^{\text{listing}} = C_{\text{total}}^{\text{listing}} \cdot \left(\frac{N_{\text{treated}}^{(All)}}{N_{\text{eligible}} \cdot p_{\text{survey}}} \right)$$

Under scale, we assume that the government would not need to conduct a separate listing survey. Instead, we proxy targeting with one month of administrative staff time per year, over five years. The discounted present value of that time is:

$$C_{\text{scale}}^{\text{listing}} = W \cdot \sum_{t=0}^{4} \frac{1}{(1+\delta)^t}$$

Administrative Costs: Administrative costs cover time spent processing scholarships. In both the pilot and scale scenarios, we assume 3 months of admin time per year for 5 years. The total discounted cost is:

$$C^{\text{admin}} = W \cdot 3 \cdot \sum_{t=0}^{4} \frac{1}{(1+\delta)^t}$$

Total Indirect Costs per Recipient:

We compute the per-recipient indirect costs as:

- **Pilot scenario:**

$$c_{\text{pilot}}^{\text{indirect}}(p) = \frac{C_{\text{pilot}}^{\text{listing}} + C^{\text{admin}}}{N_{\text{treated}}^{(p)}}$$

- **Scale scenario:**

$$c_{\text{scale}}^{\text{indirect}}(p) = \frac{C_{\text{scale}}^{\text{listing}} + C^{\text{admin}}}{N_{\text{treated}}^{(p)}}$$

These per-recipient costs feed into our program cost estimates and are reflected in the MVPF calculations. As the number of recipients differs across program variants (e.g., all students vs. girls only), the per-recipient indirect cost varies with p.

Opportunity Cost of Schooling

We estimate the opportunity cost of additional schooling by computing the present value of reduced earnings during the years of attending secondary school, from 2009 to 2012. These reductions are attributed to time out of the labor market due to continued education among treated students. We assume monthly earnings impacts are in 2020 USD and are indexed by program type p.

$$\text{Opportunity Cost}_p = \sum_{y=2009}^{2012} \frac{M_y \cdot \beta_{\text{Monthly earnings 2009-13}}(p)}{(1+0.05)^{t_y}}$$

Where:

- $\beta_{\text{Monthly earnings 2009-13}}(p)$ is the estimated treatment effect on average monthly earnings from 2009-2013 for program type p,
- M_y is the number of months during year y in which earnings are affected (12 months for all years except $M_{2012} = 6$),
- t_y is the number of years from 2008 to year y, used to discount earnings to the program start year (i.e., $t_{2009} = 1$, ..., $t_{2012} = 4.5$).

Government education expenditure

Notation

- E_y^{sec}: Government expenditure per secondary school student in year y (in 2020 USD)
- E_y^{ter}: Government expenditure per tertiary student in year y (in 2020 USD)
- $\beta_y^{\text{sec}}(p)$: Estimated increase in years of secondary education for year y of schooling under program type p
- $\beta_y^{\text{ter}}(p)$: Estimated increase in enrollment in year y of tertiary education under program type p

Secondary Education Costs

The scholarship program begins in the 2008–09 school year (Grade 10). The control group had completed on average 11.056 years of schooling for girls and 11.806 for boys. Therefore, the marginal years of education due to the program occur primarily in Grade 12 (2010–11 school year) and Grade 13 (2011–12 school year).

The government bears the cost of additional public schooling induced by the program. These costs are discounted based on when the expenditure occurs.

$$\text{GovEd}^{\text{sec}}(p) = \beta_{11}^{\text{sec}}(p) \cdot E_{2010}^{\text{sec}}(1+r)^2 + \beta_{12}^{\text{sec}}(p) \cdot E_{2011}^{\text{sec}}(1+r)^3$$

- $\beta_{11}^{\text{sec}}(p)$ is the portion of the secondary schooling increase that occurred in Grade 12 (school year 2010–11), discounted by 2 years from 2008.
- $\beta_{12}^{\text{sec}}(p)$ is the portion that occurred in Grade 13 (school year 2011–12), discounted by 3 years.
- These effects are multiplied by per-student government spending.

Tertiary Education Costs

Tertiary education impacts begin several years after the program starts. Enrollment data suggest that most students began tertiary education between 2015 and 2020, with roughly 1/3 of those who ever enrolled starting in each of three periods:

- Batch 1: 2015–2018
- Batch 2: 2016–2019
- Batch 3: 2017–2020

For each batch, we assume:

- Students attend up to 4 years of tertiary education
- Enrollment follows a linear trajectory: $\beta_1^{\text{ter}}(p)$ in year 1, increasing linearly to $\beta_4^{\text{ter}}(p)$ in year 4
- Costs are calculated using per-student expenditure from 2014 (held constant for 2015–2020) and discounted accordingly

Define slope of attendance effect:

$$\Delta_{\text{slope}}(p) = \frac{\beta_4^{\text{ter}}(p) - \beta_1^{\text{ter}}(p)}{3}$$

$$\beta_2^{\text{ter}}(p) = \beta_1^{\text{ter}}(p) + \Delta_{\text{slope}}(p) \quad ; \quad \beta_3^{\text{ter}}(p) = \beta_2^{\text{ter}}(p) + \Delta_{\text{slope}}(p)$$

$$\text{GovEd}^{\text{ter}}(p) = \sum_{b=1}^{3} \sum_{y=1}^{4} \left(\frac{1}{3} \cdot \frac{\beta_y^{\text{ter}}(p) \cdot E_{2014+y-1}^{\text{ter}}}{(1+0.05)^{d_b}} \right)$$

Discount factors by batch:

- $d_1 = 6$ (batch starting in 2015)
- $d_2 = 7$ (batch starting in 2016)
- $d_3 = 8$ (batch starting in 2017)

Explanation:

- Each batch contributes one-third of the students.
- Year-specific enrollment effects are applied per cohort year, discounted appropriately based on when costs occur relative to program start.

Total Cost

$$\text{GovEd}(p) = \text{GovEd}^{\text{sec}}(p) + \text{GovEd}^{\text{ter}}(p)$$

F.3 Estimates

Cost-Effectiveness

We calculate cost-effectiveness (CE) for second-generation cognitive development and child mortality outcomes under different program assignment and fertility scenarios.

Let:

- $\beta_{\text{cog}}(p)$: treatment effect on second generation cognitive ability (standard deviation units) for program p at age 7.

We define two fertility assumptions:

- "Not full fertility": uses actual estimated fertility per recipient.
- "Full fertility": assumes each recipient has 2.1 children (replacement fertility).

Then the cost-effectiveness for cognitive gains is calculated as:

$$\mathrm{CE}^{\mathrm{cog}}_{p,f} = \frac{\text{Scholarship Cost}_p + c^{\text{indirect}}_{\text{scale}}(p)}{n_{p,f} \cdot \beta_{\text{cog}}(p)}$$

and for mortality reduction:

$$\mathrm{CE}^{\mathrm{mort}}_{p,f} = \frac{\text{Scholarship Cost}_p + c^{\text{indirect}}_{\text{scale}}(p)}{n_{p,f} \cdot \beta_{\text{surv}}(p)}$$

Where:

$$n_{p,f} = \begin{cases} N_{\text{kids},p} & \text{if } f = \text{not-full} \\ 2.1 & \text{if } f = \text{full} \end{cases}$$

Note that the definition of costs here only includes the direct and indirect program costs, excluding any government or beneficiary externalities.

Marginal Value of Public Funds (MVPF)

We define the Marginal Value of Public Funds (MVPF) as the ratio of recipients' total willingness to pay (WTP) to the net fiscal cost borne by the government:

$$\mathrm{MVPF}^{v}_{p,s,f} = \frac{\mathrm{WTP}^{v}_{p,s,f}}{\mathrm{NetGovCost}_p}$$

where:

- p: program variant (e.g., girls-only, all),
- $s \in \{\text{low}, \text{medium}, \text{high}\}$: VSL scenario,
- $f \in \{\text{not-full}, \text{full}\}$: fertility scenario,
- $v \in \{\text{full}, \text{second-generation only}\}$: valuation scope (includes or excludes recipient earnings).

Willingness to Pay (WTP).
We compute WTP under two valuation scopes. The *full* scope includes recipient earnings gains, while the *second-generation only* scope excludes them.

$$\begin{aligned} \mathrm{WTP}^{\text{full}}_{p,s,f} &= \text{Program WTP}_p + \text{Total Survival Benefits}_{p,s,f} + \text{Lifetime Earnings Impact}_p \\ \mathrm{WTP}^{\text{second-gen only}}_{p,s,f} &= \text{Program WTP}_p + \text{Total Survival Benefits}_{p,s,f} \end{aligned}$$

Net Government Cost.
Net cost is calculated as the sum of all public expenditures minus the tax revenues recovered through income tax and VAT. This expression is independent of s, f, and v, but indexed by program p:

$$\text{NetGovCost}_p = \underbrace{\text{Scholarship Cost}_p + c^{\text{indirect}}_{\text{scale}}(p) + \text{Opportunity Cost}_p + \text{GovEd}_p}_{\text{Gross Program Cost}}$$
$$- \underbrace{\text{VAT Offset}_p + \text{VAT Revenue from Lifetime Consumption}_p + \text{Income Tax Revenue Impact}_p}_{\text{Fiscal Externalities}}$$

where:

- $c^{\text{indirect}}_{\text{scale}}(p)$: indirect cost per recipient under scale implementation,
- $\text{GovEd}_p = \text{GovEd}^{\text{sec}}(p) + \text{GovEd}^{\text{ter}}(p)$: government education costs due to increased secondary and tertiary enrollment,
- $\text{VAT Offset}_p = \tau_{\text{VAT}} \cdot \theta_{\text{VAT}} \cdot \text{Scholarship Cost}_p$,
- $\text{VAT Revenue from Lifetime Consumption}_p = \tau_{\text{VAT}} \cdot \theta_{\text{VAT}} \cdot \text{Lifetime Earnings Impact}_p$,
- $\text{Lifetime Income Tax Revenue Impact}_p$: as defined in the income tax section.

Together, these components yield the MVPF indexed by program (rows), scenario (columns), valuation scope (panels), and fertility assumption (panels), and are reported in Table 7.

Cost-Benefit Ratios

We define the Cost-Benefit Ratio (CBR) as the ratio of total government and opportunity costs to the total social benefits generated by the program. These benefits include gains to scholarship recipients and their children, as well as fiscal revenues returned to the government.

$$\text{CBR}^v_{p,s,f} = \frac{\text{Total Cost}_p}{\text{Benefits}^v_{p,s,f}}$$

where:

- p: program variant (e.g., girls-only, all),
- $s \in \{\text{low}, \text{medium}, \text{high}\}$: VSL scenario,
- $f \in \{\text{not-full}, \text{full}\}$: fertility scenario,
- $v \in \{\text{full}, \text{second-generation only}\}$: valuation scope.

Total Cost.
Costs include all government expenditures and opportunity costs of time in school:

$$\text{Total Cost}_p = \text{Scholarship Cost}_p + c^{\text{indirect}}_{\text{scale}}(p) + \text{Opportunity Cost}_p + \text{GovEd}_p$$

Total Benefits.
Benefits consist of private and fiscal gains from the program. These vary by valuation scope:

$$\begin{aligned}\text{Benefits}^{\text{full}}_{p,s,f} &= \text{Total Survival Benefits}_{p,s,f} + \text{Lifetime Earnings Impact}_p \\ &+ \text{VAT Offset}_p + \text{VAT Revenue from Lifetime Consumption}_p \\ &+ \text{Income Tax Revenue Impact}_p\end{aligned}$$

$$\begin{aligned}\text{Benefits}^{\text{second-gen only}}_{p,s,f} &= \text{Total Survival Benefits}_{p,s,f} \\ &+ \text{VAT Offset}_p + \text{VAT Revenue from Lifetime Consumption}_p \\ &+ \text{Income Tax Revenue Impact}_p\end{aligned}$$

This approach treats fiscal externalities as genuine societal benefits rather than cost offsets, consistent with a social planner perspective.

As with MVPF, these CBR estimates are reported by program variant, scenario, fertility assumption, and valuation scope in Table 7.

References

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "The Impact of Secondary School Subsidies on Career Trajectories in a Dual Labor Market: Experimental Evidence from Ghana" *Working Paper*, 2025.

Finkelstein, Amy and Nathaniel Hendren, "Welfare analysis meets causal inference," *Journal of Economic Perspectives*, 2020, 34 (4), 146–167.

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

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