Girls Outperforming Boys: Pre-school Gender Gap in Literacy and Numeracy Skills in Sub-Saharan Africa

Julie Buhl-Wiggers, Copenhagen Business School
Rebecca Thornton, Baylor University
Irina Valenzuela, University of Illinois Urbana-Champaign

Abstract

A growing body of research shows that girls of recent generations outperform boys in learning outcomes across the world. This literature focuses mainly on school-age children, with limited evidence on gender disparities at preschool ages. We use data from the Demographic and Health Surveys and Multiple Indicator Cluster Surveys across 25 Sub-Saharan African countries from 2011 to 2019 and compare mothers' reports of preschool readiness skills of boys and girls ages three and four. We find that girls are 2 percentage points more likely than boys to identify ten letters of the alphabet and to identify the first ten numbers, while we find no difference across gender in reading four words. Approximately 70 percent of the countries in our sample exhibit a female advantage in skills. We also find that girls are more likely to attend early childhood education programs and be engaged in parent-child interactions at home. We explore mechanisms and examine interactions between child gender and mother and household characteristics, as well as by country-level prevalence of early childhood education participation.

1. Introduction

Recent studies have found significant gender disparities in educational outcomes, with boys generally performing worse than girls, especially in reading (Welmond & Gregory, 2021). While the gender gap in educational achievement has long been acknowledged in high-income countries, it is increasingly common in middle-income and less developed countries (Welmond & Gregory, 2021).[1] Many argue that disparities in school-age or later-life academic achievement can be attributed to early childhood, as human capital accumulation tends to persist over time (Autor et al., 2019; DiPrete & Jennings, 2012; García et al., 2020; Heckman et al., 2013). Fundamental cognitive skills are established before the age of five; as a result, numeracy and literacy skills prior to starting school are strong indicators of performance in these subjects during fourth grade (Mullis et al., 2012; Soto-Calvo & Sánchez-Barrioluengo, 2016), and of future academic success (Heckman et al., 2013; McCoy et al., 2016).

To provide insight into the educational gaps across gender during preschool years, this paper examines the disparities in literacy and numeracy competencies among children ages 3 and 4 years old across 25 countries in Sub-Saharan Africa. We use data from the Demographic and Health Survey and Multiple Indicator Cluster Surveys collected between 2011 and 2019. In these surveys, mothers of children ages three and four years old are asked whether their child can identify at least ten letters of the alphabet, read four basic words, and identify the first ten numbers. We compare proficiency on these tasks across child gender.

We first estimate the gender gap in literacy and numeracy by pooling across all countries, including country and survey-year fixed effects as well as individual and family-level controls to account for potential imbalances across boys and girls in the sample. We find that girls perform better on average in both early literacy and numeracy. While on average across all countries, approximately 16 percent of boys can recognize 10 letters and 17 percent of boys can recognize 10 numbers, girls outperform boys in identifying letters and numbers by 2 percentage points. There is no evidence of a difference across child gender in the ability to read four basic words. We also show the gender gap by country. The vast majority of countries have gender gaps that favor girls:

[1] The 2015 Programme for International Student Assessment (PISA) finds that girls out-perform boys in reading in every country, while the gender gap in science and math varies by country (World Bank, 2018), consistent with other literature (Bertocchi & Bozzano, 2020; Buhl-Wiggers et al., 2021; Dercon & Singh, 2013; Rühle, 2022; Saito, 2011).

76 percent of the countries have a point estimate in favor of girls identifying letters, 68 percent for reading four words, and 72 for identifying numbers.

Next, we investigate several potential contributors for the gender gap in early skills, specifically, influence outside of the home through early childhood education (e.g., preschool), or inside the home (e.g., from adult caregivers). In our data, 20 percent of children (16 percent of three-year-olds and 25 percent of four-year-olds) attend an early childhood education program and over half (65 percent) have experienced at least one home stimulation activity from a caregiver (such as reading or playing with the child). We find that girls are more likely to attend early childhood education programs than boys. Moreover, parents are somewhat more likely to engage in home stimulation activities with their daughters than sons, although the magnitude of the differences are quite small.

Lastly, we examine the gender gap in outcomes by mother and household characteristics. We find that mother's literacy has a positive influence on her children's – especially daughter's – skills. In addition, although 24 percent of children do not have a father present in the household, there is no added impact of father's absence on early educational skills nor a differentiate impact across child gender. In contrast, poverty has a strong negative relationship with children's early literacy and numeracy skills, but its detrimental impact for girls is weak. Also, children living in urban areas have higher levels of pre-school skills, but not differentially across child gender.

Finally, we explore what appears to be the major predictor of girls' advantage in early literacy and numeracy: participation in early childhood education (ECE). We estimate the gender gap in early childhood skills by comparing countries with low overall ECE attendance to those with high ECE attendance. We find that the gender gap favoring girls is stronger in countries with high levels of ECE participation, and nearly disappears in countries with low rates of ECE participation.

We make several contributions to the existing literature. First, our paper adds to the expanding body of literature that documents boys lagging behind girls across the world (Bando et al., 2024; Buhl-Wiggers et al., 2021; Lai, 2010; Nakajima et al., 2016; Spaull & Makaluza, 2019; Suryadarma, 2015). Most prior research in developing countries has concentrated on school-aged boys and girls, whereas our paper is among the few that assess early numeracy and literacy skills. Two exceptions are Bago et al. (2019) and Bando et al. (2024), that examines the early childhood gap in developing countries. However, the former only examined one country, Ghana, and a multidimensional index of six childhood abilities, including literacy and numeracy. While their

findings revealed no evidence of gender differences in children's development in Ghana, their study was relatively underpowered. The latter study focused on children aged 7 to 48 months from nine countries, primarily in South America and found that girls surpass boys in language and socioemotional skills. In contrast, our study focuses on numeracy and literacy abilities, using pooled data from 25 countries to have a larger sample– the sample size in our study is important for power since the magnitude of the gender gap is relatively small.

Second, our research contributes to the existing debate on the determinants of the educational gender gap. On the one hand, there has been an extensive discussion regarding the gender gap in mathematics and STEM-related fields, which favors boys over girls. These factors include biological differences (Wilder & Powell, 1989), the influence of role models (Kahn & Ginther, 2017), and perceptions about one's math ability (Bharadwaj et al., 2016). On the other hand, a range of factors have been proposed to explain gender differences in educational outcomes more broadly, including labor market (Heath & Mobarak, 2015), income (Björkman-Nyqvist, 2013), and cultural differences (Dickerson et al., 2015). Our paper contributes to this literature by highlighting the significance of early childhood education and child-parents interaction in home stimulation activities as the potential predictors of early literacy/numeracy skills.

Third, this paper contributes to the existing literature on the intergenerational effects of parents' educational backgrounds on their children's outcomes. We contribute to this strand of literature by presenting empirical findings that establish a correlation between early childhood literacy skills and the literacy attainment of parents. We find that mothers' literacy appears to be of greater significance for their daughters.

Fourth, fathers have been found to be very important for children's outcomes, especially for boys (Diniz et al., 2021; Henry et al., 2020; S. Lundberg, 2022; Reeves, 2022; Wasserman, 2020). Almost all of the existing literature on the role of fathers comes from high income settings. We provide some of the first analysis of the impact of a father's absence on educational performance in a low-income setting. We find no difference in the likelihood of having a father present by gender, and contrary to the existing literature in higher-income settings, find that fathers do not differentially contribute to their pre-school son's vs daughter's early grade skills.

The remainder of the paper is organized as follows. Section 2 describes the data and summary statistics. Section 3 presents the empirical approach. Section 4 presents the results. Section 5 provides a discussion of the results while Section 6 concludes.

2. Previous Literature

2.1 Gender Gaps in Developing Countries

While the gender gap favoring girls in educational outcomes in higher income countries has been well-documented, there is also increasing evidence for the same tendency in developing countries, especially in reading (Grant & Behrman, 2010; Spaull & Makaluza, 2019; T. L. Zuze & Reddy, 2014). The evidence of a gender gap favoring girls in math is less clear (Bharadwaj et al., 2016; Dickerson et al., 2015; Ng'ang'a et al., 2018). Studies in Indonesia find that girls outperform boys in literary and language, as well as numeracy and math in primary school, and that the gap widens during secondary school (Nakajima et al., 2016; Suryadarma, 2015). In China, Lai (2010) finds that girls outperform boys in math and science from primary through middle school, while in contrast girls are found to under-perform boys in math tests in Chile (Bharadwaj et al., 2016). Research in South Africa indicates that girls perform better than boys in literacy and that the gender gap in math has been shrinking or disappearing in recent years (Spaull & Makaluza, 2019; L. Zuze et al., 2017), while Rühle (2022) finds that girls outperform boys in math and science in grade five, diminishing in higher grades. Evidence from Uganda, Kenya, and Tanzania finds that girls perform better in literacy and numeracy than boys of primary school ages, but that this gender gap varies greatly within countries (Buhl-Wiggers et al., 2021).

2.2 Mechanisms

There are several hypothesized mechanisms as to the determinants of boys falling behind girls in education. Some studies have linked the lower levels of schooling among boys to their higher likelihood of social and behavioral problems in school, including difficulties with self-regulation and attention (Buchmann et al., 2008; Owens, 2016). Boys are more likely to be expelled from pre-kindergarten and are more likely to be disciplined or suspended than girls in early elementary school, persisting through high school (DiPrete & Jennings, 2012; Downey & Vogt Yuan, 2005; Gilliam & others, 2005).

Family structure is another potential factor, with studies suggesting that the absence of male role models in single-parent households leading to underachievement and school problems among boys (Autor & Wasserman, 2013; Lundberg, 2022; Welmond & Gregory, 2021). Gerrand & Nduna (2021) find that South African girls can find strength even with the loss of a father figure. Older papers have examined the role of mother's and father's separate influence on their children's

schooling in Africa (Lloyd & Blanc, 1996), while Madhavan et al. (2017) find that boys and girls in South Africa are equally affected with non-nuclear household structures while boys may be more responsive to having another non-parental family member. Much of this literature focuses on family structure affected by the HIV/AIDS crisis in Africa (Timaeus & Boler, 2007).

Some studies have found that family disadvantage has a more detrimental effect on boys than girls. For example, girls from low- and middle-socioeconomic status (SES) families have an advantage in reading, while boys from high-SES families have an advantage in numeracy (Autor et al., 2023; Cobb-Clark & Moschion, 2017). Brenøe & Lundberg (2018) find, using linked administrative Danish data, that maternal education has a more significant impact on the education and employment outcomes of daughters than sons, with effect remaining stable over time. In contrast, they found that paternal education has a smaller effect on the gender education gap, which favors sons. Akresh et al. (2023) demonstrate in Indonesia that a mother's higher education, obtained through exposure to a school construction program, leads to increased secondary and tertiary schooling for her children. In Africa, Glick & Sahn (2000) and Buhl-Wiggers et al. (2021) find that maternal education significantly impacts daughters' education. Supporting the idea that boys may be more sensitive to disadvantage, recent evidence points to a more pronounced gender gap in academic outcomes at the distribution tails of outcomes, where there is an overrepresentation of boys (Autor et al., 2023).

Parents play a fundamental role as primary caregivers in a child's early years. The quality of parent-child relationships and parental support for learning during this crucial period greatly influence early child development and learning outcomes (Anders et al., 2013; Jeong et al., 2021). For example, Jeong et al. (2021) found that interventions targeting parenting knowledge, parenting practices, and parent-child interactions improved cognitive, language, and socio-emotional development in young children. During the early years, the home environment is a central context for learning and development (McCoy et al., 2018); it provides a conducive space for children to engage in literacy and numeracy activities (Soto-Calvo & Sánchez-Barrioluengo, 2016). This motivates us to examine the gender gap in home stimulation activities.

In this paper, we also examine the possible role of early education. In recent times, there has been a global expansion in out-of-home early educational services for children. As of 2015, approximately 40 countries, including several lower- and middle-income countries, have mandated universal access to preprimary education (UNESCO, 2015). Studies have demonstrated that

participation in early childhood education programs serves as a reliable predictor of a child's early numeracy skills (Anders et al., 2013). Additionally, preschool attendance has been found to be associated with improved academic performance and longer-term benefits for individuals throughout their lives (García et al., 2021). In the United States, the Perry Preschool Project, a high-quality early childhood education program that targeted disadvantaged African American children has been found to have long term effects on cognitive abilities, marital stability, earnings, criminal behavior, and health (García et al., 2021), and that the positive impact of this preschool program extended to the subsequent generation, as the children of the original participants exhibited higher educational attainment, employment rates, reduced involvement in criminal activities, and better health compared to the children in the control group (García et al., 2021). Most notably, papers from the United States have found that participation in preschool programs, while beneficial for both genders, are particularly beneficial for boys for outcomes such as special education and grade retention (Magnuson et al., 2016).

3. Data

We use two rounds of the Demographic and Health Survey (DHS) and Multiple Indicator Cluster Surveys (MICS) from 25 African countries. The DHS and MICS are large-scale, nationally representative household surveys.[2] Appendix Table A2 presents a summary of detailed variable descriptions. We use data collected from mothers of children ages three and four.[3]

3.1 Variables

The Early Childhood Development Index (ECDI) questionnaire consists of ten caregiver-reported questions designed for children ages three to four years old across four domains of development: literacy-numeracy, learning/cognition, physical development, and socio emotional

[2] USAID has supported the DHS program since its establishment in 1984 while MICS was created by UNICEF in 1995. DHS and MICS employ a comparable sampling design and are highly comparable data sources (Hancioglu & Arnold, 2013). Data are collected through standardized, face-to-face interviews in low- and middle-income countries.
[3] In DHS surveys, information on children under the age of five is collected from biological mothers with no information on children who are orphaned or not living with biological mothers. In MICS surveys, information on all children under five is collected from mothers or primary caregivers in the household, regardless of whether their biological mothers reside in the same household and include orphans and foster children (Hancioglu & Arnold, 2013). For this paper, we exclusively utilize data on children whose mothers have been interviewed, allowing us to merge information about the children with that of their mothers.

development (Loizillon et al., 2017). We focus on three yes/no questions reflecting early literacy and numeracy skills to proxy for literacy and numeracy pre-school skills: 1) "Can your child name or identify at least ten letters of the alphabet?"; 2) "Can your child read at least four simple, popular words?"; and 3) "Can your child name and recognize the symbols of all numbers from 1-10?." These proxies for early literacy and numeracy are based on mother' self-reports which is less desirable as an outcome as compared to objective literacy and numeracy assessments. To the extent that any bias in self-reports is uncorrelated with child gender, our analyses still provide an unbiased estimate of the gender gap. In addition, we also explore the potential for self-reporting bias by presenting results separately by mother's literacy.

To assess whether participation in early childhood education contributes to early gender disparities, we utilize a question that is available in both MICS and DHS. This question asks each mother to report whether her 3-4-year-old child attends an early childhood education program.

To evaluate the role of home learning activities in explaining early gender disparities, we utilize questions from MICS and DHS that inquire about six basic activities undertaken by caregivers with their children in the three days before the survey. These home stimulation activities include reading books, telling stories, singing songs, taking the child outside, playing with the child, and engaging in naming, counting, or drawing activities.[4] In addition, mothers were asked to specify which family member (mother, father, or another relative) participated in these activities with the child. We create an overall measure of parental home stimulation activities, reflecting the total number of activities children engage in with at least one of their parents. We also create binary indicator variables separately for each activity. Lastly, we create binary indicators separately by each parent providing the engagement. For example, we create separate binary indicator variables for caregiver participation, indicating if an activity was carried out by only the mother, only the father, both parents, or only another family member.

We use six variables to control for mother and household characteristics including: mother's age, an indicator for mother's completed secondary education level or higher, an indicator that the mother is literate, number of household members, an indicator of the child's father not being part

[4] Several of these six home stimulation activities are directly related to literacy and numeracy skills. Reading and storytelling can indicate engagement in literacy-related activities that foster early vocabulary development, comprehension, and reasoning. Naming, counting, and drawing activities encourage verbal interaction between caregivers and children, enhancing vocabulary, math skills, and fine motor development (McCoy et al., 2016). Playing, singing songs, and exploring the outdoors provide cognitive stimulation, model positive social and emotional behaviors, and facilitate learning interactions with the physical and social environment (McCoy et al., 2016).

of the household, an indicator of being poor (as measured by being in the lowest two wealth quintiles), and an indicator of living in an urban area.

3.2 Sample

We use the surveys collected in Sub-Saharan African countries that had available data about a child's early literacy and numeracy skills, early childhood education attendance, and home stimulation activities.[5] Our final data consists of rounds six and seven from the DHS from eight countries and rounds five and six from the MICS from 23 countries, collected between 2011 and 2019. Appendix Table A1 provides the list of the 25 countries and the years of survey participation.

We exclude any observations with missing values in any of the three learning outcomes, early childhood education attendance, home stimulation activities, mother's literacy, education level, or age, resulting in an analytical sample of 136,531 children.[6]

3.3 Summary Statistics

Table 1 presents the summary statistics of the children in our analytical sample. Panel A presents the statistics on children's outcomes. Boys (N=68,510) and girls (N=68,021) are equally represented in the sample. The average age in the sample is 3.5 years old and is even across gender. On average, 17 percent of children aged 3-4 in our sample can name/identify ten letters of the alphabet, 13 percent can read four words, and 18 percent can identify the first ten numbers. Only 6 percent of children are proficient in all three skills.[7] Across the 25 countries, 20 percent of pre-schooled-age children attend an early childhood education program.[8]

Table 1, Panel B depicts the summary statistics of parental stimulation activities at home. There are six home stimulation activities in which at least one parent (with or without the participation of a close relative) interact with their child. The two most popular activities are "sing songs to

[5] We exclude Sierra Leone because it lacks information about early childhood education.
[6] Our analytical sample loses a total of 27,535 observations (17 percent). The three countries with the highest number of missing learning outcomes include: Cameroon, Chad, and Democratic Republic of the Congo, each from DHS surveys that restricted early childhood development related questions to households selected for hemoglobin measurements (Cameroon, DHS 2011 and DR Congo, DHS 2013), or in households not selected for the men's survey (Chad, DHS 2014).
[7] The percentage of boys who can identify the ten letters of the alphabet is 16 percent, compared to 18 percent of girls (Table 1, Panel A, Columns 3 and 5). There are similar differences in identifying the first ten numbers, where 17 percent of boys have this ability compared to 19 percent of girls. The ability to read four basic words has the smallest difference between genders: 12 percent of boys vs. 13 percent of girls.
[8] In the sample, 19.5 percent of boys and 21 percent of girls attend an early childhood education program.

child" (40 percent) and "taking the child outside" (44 percent); 65 percent of children experience at least one home-based parental stimulation activity and the average child experiences between one and two (1.8) different home stimulation activities from at least one of their parents.[9] Mothers conduct the majority of the home stimulation activities (1.2 activities), followed by other relatives alone (1.2 activities), both parents together (0.4 activities), and lastly, fathers (0.2 activities).[10]

Table 1, Panel C, presents summary statistics for some basic household and mother characteristics in the sample. On average, mothers are 31 years old. Less than 25 percent of mothers in the sample have completed at least secondary education or possess a higher level of education.[11] Less than half, 37 percent, of mothers in the sample report being literate. In almost one out of four households in our sample, the father is not present, which could be due to death, divorce, or separation. We create a dummy variable for being poor which takes the value one if the wealth status of the household is 1 or 2 out of 5, and zero otherwise. Almost half of the households in our sample are characterized as poor according to this definition. Most of the households in the sample are located in rural areas (72 percent).

There are some statistically significant differences between boys and girls in some household characteristics – boys are more likely to have less educated mothers and more likely to come from poor households. These differences motivate us to control for these household characteristics in all regressions.

4. Empirical Strategy

This section describes how we estimate gender differences, using ordinary least squares (OLS) regressions with country and survey-year fixed effects. Our preferred estimate pools across all countries and survey waves and compares the average of each of the three literacy/numeracy indicators across gender. Pooling across countries is important for statistical power.

[9] For most home stimulation activities, girls are involved more frequently than boys. The only activity where boys participate more is "taking the child outside", the difference is not statistically significant (Columns 3 and 6).

[10] Girls are equally likely as boys to participate in home stimulation activities from their mother alone. Boys are more likely to participate in activities from their father alone.

[11] In the MICS and DHS surveys, the mother's education variable is measured as a 4-level categorical education variable that is harmonized across countries: 1 pre-primary or none, 2 primary, 3 secondary, and 4 higher. There is no information about the number of years of education.

Table 1. Summary Statistics

	All		Boys		Girls		Diff
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	(Boys - Girls)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Children Variables							
Identify 10 letters	0.169	0.375	0.160	0.367	0.178	0.382	-0.02***
Read 4 words	0.125	0.331	0.122	0.327	0.129	0.335	-0.01***
Identify 10 numbers	0.177	0.382	0.169	0.375	0.186	0.389	-0.02***
Have all 3 skills	0.061	0.239	0.057	0.232	0.065	0.246	-0.01***
Attends Early Childhood Education	0.204	0.403	0.196	0.397	0.211	0.408	-0.02***
Panel B. Parental Home Stimulation (HS) Activities							
Read to child	0.147	0.354	0.145	0.352	0.148	0.355	-0.00
Tell stories to child	0.317	0.465	0.313	0.464	0.322	0.467	-0.01***
Sing songs to child	0.399	0.490	0.390	0.488	0.407	0.491	-0.02***
Take child outside home	0.439	0.496	0.441	0.496	0.437	0.496	0.00
Play with child	0.315	0.465	0.314	0.464	0.317	0.465	-0.00
Name/count/draw with child	0.228	0.420	0.226	0.418	0.230	0.421	-0.00
At least one home stimulation activity	0.654	0.476	0.652	0.476	0.656	0.475	-0.00
N of Parental home stimulation activities	1.844	1.845	1.829	1.837	1.860	1.854	-0.03**
N of HS conducted by mothers	1.227	1.522	1.173	1.495	1.281	1.546	-0.11***
N of HS conducted by fathers	0.227	0.669	0.254	0.702	0.199	0.634	0.06***
N of HS conducted by both parents	0.391	0.955	0.401	0.966	0.380	0.943	0.02***
N of HS conducted by other relatives	1.178	1.541	1.174	1.534	1.181	1.549	-0.01
Panel C. Mother and Household Characteristics							
Mother's Age	30.795	6.891	30.769	6.874	30.822	6.908	-0.05
Mother secondary education or higher	0.241	0.428	0.239	0.427	0.243	0.429	-0.00
Literate mother	0.367	0.482	0.365	0.481	0.369	0.482	-0.00
HH members	7.877	5.202	7.889	5.203	7.865	5.201	0.02
Father is not part of the household	0.239	0.426	0.234	0.424	0.243	0.429	-0.01***
Poor	0.477	0.499	0.480	0.500	0.474	0.499	0.01*
Urban	0.279	0.448	0.278	0.448	0.280	0.449	-0.00
Observations	136,531		68,510		68,021		136,531

Source: Author's calculation using DHS (wave 6 and 7) and MICS (wave 5 and 6) for selected Sub-Saharan African countries.

Notes: This table shows information from the children's questionnaire and the mother's questionnaire. Observations were dropped if any of the children's learning outcomes and/or the mother's literacy and/or early childhood education attendance and/or home stimulation activities were missing. The average number of home stimulation activities is measured amongst all children. If we focus on those children who experience at least one home stimulation activity the average increases to 2.8. Columns 1, 3, and 5 show the mean values of each variable. Columns 2, 4, and 6 present the standard deviation. Column 7 shows the average difference between boys and girls. ***, **, and * denotes statistical difference at the 1, 5, and 10%-level, respectively.

We estimate the following OLS regression:

$$Y_{ic} = \beta_0 + \beta_1 Female_{ic} + \beta_2 X_{ic} + \gamma_c + \gamma_y + e_{ic} \quad (1)$$

Y_{ic} is an outcome variable for child i in country c. Our main school readiness variables are indicators if the child can 1) identify ten letters of the alphabet, 2) read four simple words, and 3) identify the first ten numbers. In addition, we also measure if the child is attending an early childhood education program, indicators of receiving home stimulation activities, and an indicator of receiving any home stimulation activities. We also investigate if home stimulation activities differ by the parent who engages in the activity, by separately estimating Equation (2) by type of home stimulation activity and by the parent engaging in each activity.

$Female_{ic}$ is a dummy variable that indicates when the child is female. To control for the differences in the descriptive statistics of household characteristics across boys and girls, we include a set of household controls, X_{ic}, which include: children's age, mother's age, an indicator of mother's having secondary or higher education level, number of household members, an indicator of father's absence in the household, an indicator of household being poor, and an indicator of urban residence. γ_c and γ_y indicate country and survey-year fixed effects that enter into the model as indicator variables.[12] Robust standard errors are clustered at the country level. In all cases, we use sampling weights provided by the MICS and DHS for nationally representative estimates. The estimated coefficient, β_1, represents the gender difference in the educational skills.

While Equation (1) provides our preferred estimates of the gender gaps using data pooled across all available countries, to understand the variation in the gender gaps across country, we also estimate Equation (1) separately for each of the 25 countries in the sample. For these estimates, we only include survey-year fixed effects, omitting country fixed effects, and we visually plot the coefficients, β_1, for each country.

To broaden our understanding of potential mechanisms underlying the gender gap in skills we estimate heterogeneity of the gender gap in literacy/numeracy skills by mother's literacy, father's

[12] For robustness, we also include village fixed effects, and the results barely change (see Appendix Table A3). Unfortunately, we cannot include household fixed effects as the sample reduces sharply when restricting to households with both a boy and a girl.

absence in the household, indicators of poverty and urban residence. To do so, we estimate the following OLS model:

$$Y_{ic} = \alpha_0 + \alpha_1 Female_{ic} + \alpha_2 Z_{ic} + \alpha_3 Female_{ic} * Z_{ic} + \alpha_4 X_{ic} + \gamma_c + \gamma_y + e_{ic} \quad (2)$$

where, Z_{ic} includes indicators for mothers' literacy, father's absence, living in a poor household, and living in an urban area. We run these in separate models with a single interaction term, as well as a fully interacted model with all four interaction variables.

In addition to the mother/household-level interaction terms, we also generate a country-level indicator of having above median level of early childhood education attendance. Appendix Figure A1 shows the proportion of early childhood education attendance by country. We construct a zero one variable if the country average is above or below the median rate of attendance and include this in an interacted model, following Equation (2).

5. Results

5.1 Gender Differences in Childhood Literacy and Numeracy Skills

This section presents the estimates of gender disparities in children ages 3 to 4 years on early educational skills using all available data of Sub-Saharan African countries from 2011 to 2019. Table 2 show the estimates of the gender gap from Equation (1), which includes controls for children, mothers, and household characteristics as well as country and survey-year fixed effects.[13] On average, girls are 1.6 percentage points more likely to identify ten letters of the alphabet and 1.6 percentage points more likely to identify the first ten numbers compared to boys. There is no difference in the likelihood of reading at least four words.[14]

[13] Appendix Table A4 presents the results including village fixed effects instead of country fixed effects and the results barely change.

[14] One potential concern is the bias in mothers' self-reporting, particularly since illiterate mothers may struggle to accurately evaluate their children's abilities. Table A3 in the Appendix presents results separately by mothers' literacy levels, showing that illiterate mothers tend to underestimate the gender gap in favor of girls, especially in the case of identifying 10 letters of the alphabet. If this outcome is due to illiterate mothers being less likely to recognize their children's abilities, then our estimates may represent a lower bound of the gender gap favoring girls.

Table 2. Gender Effects on Preschool Literacy/Numeracy

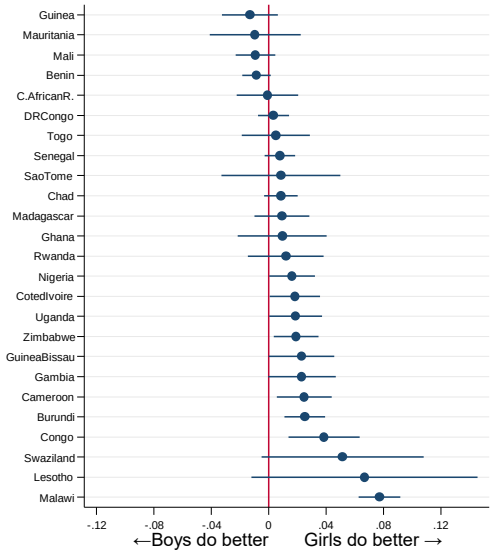
	Identify 10 letters	Read 4 words	Identify 10 numbers
	(1)	(2)	(3)
Female	0.016**	0.005	0.016***
	(0.006)	(0.004)	(0.005)
Overall mean	0.169	0.125	0.177
Observations	136,531	136,531	136,531
R-squared	0.155	0.081	0.152
Country FE	YES	YES	YES
Survey-Year FE	YES	YES	YES
Controls	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. The independent variables in Columns 1-3 are dummy variables that measure whether the child shows early literacy or numeracy skills based on their mother's response. All columns include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's absence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

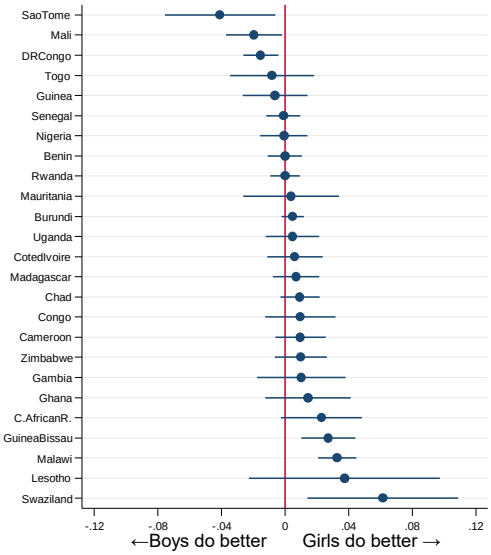
We also estimate the gender differences of early educational skills separately for each of the 25 countries in our sample to illustrate how the gender gap varies across countries and to identify in how many countries it is present. Figure 1 presents the gender gap in literacy (Panels A and B) and numeracy (Panel C) and shows that the point estimates fluctuate between -4 and 9 percentage points for the three different learning outcomes. On average, more than 15 countries (over 70 percent) report gender gaps in favor of girls.

Figure 1. Numeracy/Literacy Gender Gap by Country

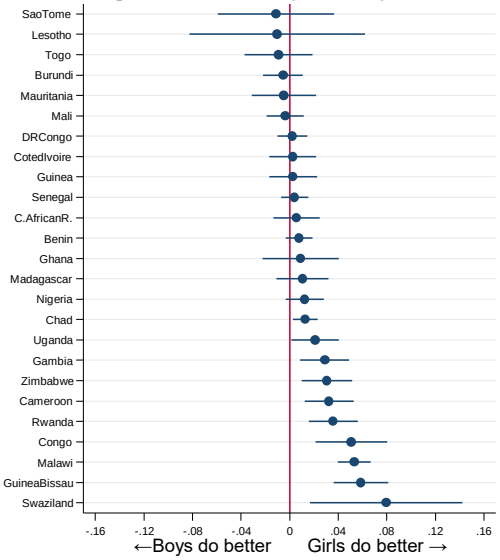
(A) Gender Gap on Child Literacy: Identify 10 Letters



(B) Gender Gap on Child Literacy: Read 4 Words



(C) Gender Gap on Child Literacy: Identify 10 Numbers



Notes: Each figure presents the estimated coefficient of the gender disparity from a separate country regression based on Equation (1). The 95 percent confidence intervals are depicted as horizontal bars. All figures include controls: child's age, mother's age, mother's having at least secondary education, household size, father's absence in household, household poverty status, and residence area (urban/rural). All regressions include population weights and survey-year fixed effects.

5.2 Gender Differences in Early Childhood Education and Home Stimulation

Table 3 presents the results on the gender difference in early childhood education attendance and parental home stimulation. Columns 1 and 2 show that girls are 1.1 percentage points more likely to attend early childhood education, and 0.6 percentage points more likely to receive any home stimulation activities from at least one parent. Columns 3 – 8 show the results for individual home stimulation activities. Girls are almost one and two percentage points more likely to have at least one parent tell stories and sing songs to them, respectively, compared to boys. Parents are more engaged in naming, counting, and drawing with girls than boys. However, the magnitudes of these differences are small, especially relative to the overall mean of these reported activities.

Table 3. Gender Effects on Early Childhood Education and Home Stimulation

	Early Childhood Education Program	Has at least one parental home stimulation activity	Parental Stimulation Activity at Home					
			Read to children	Tell stories to children	Sing songs to children	Take child outside home	Play with child	Name / Count / draw with child
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.011**	0.006*	0.002	0.009**	0.018***	-0.005	0.005	0.004**
	(0.005)	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)
Overall mean	0.204	0.654	0.147	0.317	0.399	0.439	0.315	0.228
Observations	136,531	136,531	136,531	136,531	136,531	136,531	136,531	136,531

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In Column 1, the dependent variable indicates whether the child attends any early childhood education program. In Column 2, the dependent variable equals 1 if the child receives at least one home stimulation activity from any of their parents (mother or father or both parents) and 0 if the child does not participate in any activity or if these activities are carried out by only close family members. In Columns 3 to 8, the dependent variables are indicators of whether the child receives any of the home stimulation activity from at least one parent. All estimates include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's absence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The MISC and DHS surveys allow us to investigate the gender disparity in home stimulation activities, separately by which parent is engaged in each activity. Table 4 present the results; Panel A presents the results for home stimulation activities defined as 1 if they are carried out by the mother, Panel B presents the results for activities carried out by the father, Panel C for activities carried out by both parents, and Panel D carried out by other family members.[15]

[15] Panels A, B, and C also include participation of other family members, while Panel D pertains to activities conducted solely by other family members without the involvement of any parent.

Table 4. Gender Effects on Home Stimulation Activities by Parent

	Read to children	Tell stories to children	Sing songs to children	Take child outside home	Play with child	Name / Count / draw with child
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Home stimulation activities carried out by mothers					
Female	0.003	0.015***	0.022***	0.034***	0.016***	0.008***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)
Overall mean	0.09	0.20	0.31	0.28	0.19	0.15
Observations	136,531	136,531	136,531	136,531	136,531	136,531
	Panel B: Home stimulation activities carried out by fathers					
Female	-0.002**	-0.007***	-0.002**	-0.027***	-0.009***	-0.004***
	(0.001)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
Overall mean	0.02	0.05	0.03	0.06	0.03	0.03
Observations	136,531	136,531	136,531	136,531	136,531	136,531
	Panel C: Home stimulation activities carried out by both parents					
Female	0.001	0.001	-0.002	-0.011***	-0.002	-0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Overall mean	0.03	0.07	0.06	0.10	0.09	0.04
Observations	136,531	136,531	136,531	136,531	136,531	136,531
	Panel D Home stimulation activities carried out by other family members					
Female	0.002	-0.003	0.004*	-0.001	-0.008	-0.003
	(0.002)	(0.003)	(0.002)	(0.002)	(0.006)	(0.003)
Overall mean	0.09	0.14	0.19	0.19	0.40	0.16
Observations	136,531	136,531	136,531	136,531	136,531	136,531
Country FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In Columns 1 to 6, the dependent variables are indicators of whether the child receives any of the home stimulation activity. Panel A defines the dependent variables as 1 if the activities are carried out by only the mother or the mother and another relative, and 0 otherwise. Panel B defines them as carried out by only the father or the father and another relative. Panel C defines them as carried out by at least both parents, and Panel D defines them as carried out by only other family members, excluding parents. All columns include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's absence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4, Panel A shows that, except for reading (Column 1), girls are 1 to 3 percentage points more likely to participate in each of the home stimulation activities when these activities are primarily conducted by their mothers. In contrast, Panel B reveals the opposite pattern. When the parent-child interaction mainly involves the father, boys have an advantage over girls in activities such as storytelling, playing with the children, and naming and counting, with boys being almost three percentage points more likely to engage in the activity of taking the child outside the home.

However, when both parents participate in these activities, there is no clear advantage for either girls or boys, except for boys being more likely to be taken outside by their parents. Additionally, when these activities involve only other relatives, there is no significant difference across child gender, except for a slight advantage for girls in being sung a song. Overall, it appears that there is a same-sex bias—mothers tend to engage more with their daughters, while fathers tend to engage more with their sons. Consequently, since a higher percentage of children interact in these activities exclusively conducted by their mothers, and their mothers participate in a greater number of these activities, this could disadvantage boys.

5.3 Heterogeneity: Does the Gender Gap Vary Across Characteristics?

To further understand the gender gap in literacy and numeracy skills, we examine how the gender gap varies with mother and household characteristics. Household characteristics, such as parental behavior, education and economic conditions, can significantly influence children's learning opportunities (Soto-Calvo & Sánchez-Barrioluengo, 2016). Following Equation (2), we examine interactions between gender and mother's literacy, father's absence in the household, and indicators of poverty and urban residence.

Table 5 presents the results for the fully interacted model while Appendix Tables A5 to A7 present each interaction term separately. For both boys and girls, having an illiterate mother is associated with lower literacy and numeracy skills. Our results echo previous findings for early childhood literacy and numeracy and show that literate mothers have a greater impact on their daughters' ability to identify letters and numbers, although this pattern is not present for reading four words. A father's absence from the household has no statistically significant impact on children's educational skills nor is there any relationship between a father's absence and gender across all three outcomes. Several papers have found that the gender gap in educational varies according to poverty, where girls are more disadvantaged in poorer areas and vice versa for boys

(Buhl-Wiggers et al., 2021; Dickerson et al., 2015). Poverty is associated with lower literacy and numeracy skills for both boys and girls, but we only find limited evidence for this contributing to the gender gap. Living in urban areas is associated with greater early literacy and numeracy skills for children, but it does not appear to contribute to a gender advantage in these skills.

Table 5. Heterogeneity of the Gender Gap

	Identify 10 letters	Read 4 Words	Identify 10 Numbers
	(1)	(2)	(3)
Female	0.016	0.004	0.010
	(0.010)	(0.006)	(0.008)
Literate mother	0.120***	0.083***	0.126***
	(0.024)	(0.020)	(0.026)
Father's absence in the household	-0.000	-0.006	-0.003
	(0.005)	(0.005)	(0.005)
Poor	-0.076***	-0.054***	-0.082***
	(0.016)	(0.013)	(0.015)
Urban	0.087***	0.046***	0.078***
	(0.014)	(0.010)	(0.009)
Mother's literacy x female	0.023***	0.007	0.015**
	(0.005)	(0.006)	(0.007)
Father's absence in the household x female	-0.002	0.005	0.002
	(0.008)	(0.006)	(0.006)
Poor x Female	-0.011	-0.002	-0.004
	(0.008)	(0.007)	(0.008)
Urban x Female	-0.009	-0.007	0.008
	(0.009)	(0.007)	(0.009)
R-squared	0.158	0.157	0.157
Observations	136,531	136,531	136,531
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Controls	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. All columns include all interaction terms in the respective regression. In addition to the explanatory variables and its interaction term with child's gender, all columns include the following controls: child's age, mother's age, and household size. All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Lastly, we estimate the gender gap separately by countries with high and low levels of early childhood education. High levels of ECE are defined as equal to or above the median level of early childhood education calculated from all countries in the sample, while low levels are defined as below the median. Table 6 presents shows that the gender gap almost disappears for countries with low levels of early childhood education. These results may suggest that differences in early childhood education could be a driver of gender differences in early childhood skills.

Table 6. Gender Effects on Preschool Literacy/Numeracy by Early Childhood Education (ECE) Participation Level

	Identify 10 letters	Read 4 Words	Identify 10 Numbers
	(1)	(2)	(3)
	Panel A: Countries with high ECE		
Female	0.029**	0.010*	0.026***
	(0.010)	(0.006)	(0.007)
Observations	63,622	63,622	63,622
R-squared	0.186	0.096	0.173
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Controls	YES	YES	YES
	Panel B: Countries with low ECE		
Female	0.005	0.001	0.008**
	(0.004)	(0.004)	(0.004)
Observations	72,909	72,909	72,909
R-squared	0.105	0.061	0.087
Country FE	YES	YES	YES
Survey-Year FE	YES	YES	YES
Controls	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. Panel A includes only countries with a high level of early childhood education (ECE) program participation (equal or above the median of all countries in the sample). Panel B includes countries with a low level of ECE program attendance. In columns 1, 2, and 3 the dependent variables are dummy variables that measure whether the child shows early literacy or numeracy skills based on their mother's response. All columns include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's absence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6. Discussion

This study estimates the early gender gap in developing countries and evaluates some potential the factors driving this gap.

First, we find that girls outperform boys in early literacy and numeracy skills, such as identifying ten letters of the alphabet and recognizing ten numbers, as reported by their mothers. This finding complements previous results regarding girls' advantage at the school level and provides valuable insights into the role of early abilities. Moreover, it aligns with the findings of Bando et al. (2024), who demonstrate a pro-girl trend in early childhood skills.

Second, we contribute to the evidence on the gender gap in learning using the latest information from the MICS and DHS surveys. This is particularly important for numeric skills, where findings in developing countries have been mixed. Some studies have shown that boys outperform girls in mathematics and science (Dickerson et al., 2015), while more recent studies indicate that African girls perform better than boys (Buhl-Wiggers et al., 2021; Spaull & Makaluza, 2019).

Third, we find that the gender disparity in early numeracy and literacy skills could be partially attributed to girls being more likely to attend early childhood education programs and receiving more learning stimulation activities at home than boys. Additionally, maternal literacy plays a significant role in their daughters' early literacy. All three contributing factors to the advantage of girls depend on parental involvement and behaviors related to learning activities at and away from home, and the educational context of parents. The results on early childhood stimulation suggest that parents may be more inclined to invest more in their daughters than their sons. This leads to the question of why parents show this differential investment.

Fourth, even though we find that fathers are absent from almost a quarter of the households this does not seem to play a differential role in educational skills. There are small same-sex home stimulations, with fathers interacting more with their sons, and mothers interacting more with their daughters, but this unlikely to play a large role in the gender gaps we observe.

One potential reason for disparities in parental investment is parental preference. Educated mothers, in particular, may have a stronger preference for educating their daughters and therefore invest relatively more in their education. Another potential reason is parental beliefs. Parents may invest more in their highly skilled children if they believe that such investment, particularly in parental teaching, will yield greater benefits (Chuan et al., 2022). For example, Chuan et al. (2022) provide evidence that a gap in beliefs partially drives the gender gap in parental investments. They

demonstrate that beliefs about reading and math abilities are more optimistic for daughters than for sons, and parents of girls are more confident about their daughters' likelihood of attending college compared to parents of boys. A third reason is the difference in children's behavior, which may facilitate parental investment. Parents may be more inclined to teach girls if they find it easier than teaching boys. Chuan et al. (2022) found that girls aged 3-5 scored higher than boys in self-regulation (the ability to sit still and focus) and that parents were more likely to report that girls enjoy being taught than boys. These are some potential reasons that could explain the relative advantage of girls in non-mathematical skills.

This paper has certain limitations. First, although the analysis covers a substantial number of countries in Sub-Saharan Africa, the specific context of these countries may not necessarily represent the situation in other low-income countries, such as those in Latin America. Second, the data does not allow us to compare boy and girl siblings and thus we cannot rule out that unobserved household characteristics could bias the results. Third, all child-related variables recorded in the surveys are based on reports provided by mothers, and they may be subject to recall bias and misreporting. Moreover, the outcomes reported by mostly female adults can pose some difficulties for their measurement. First, reported differences in these skills may reflect differences in social and cultural norms surrounding early childhood education between countries, rather than children's cognitive ability. Second, these literacy-numeracy questions are more sophisticated than other comparable instruments that emphasize, for instance, counting rather than symbol recognition (McCoy et al., 2016). Importantly, if there are cultural biases in favor (or against) girls, our analysis would pick up these biases rather than actual gender gaps in skills.

Third, the literacy and numeracy questions in the MICS and DHS surveys may not fully capture the developmental domains of children in this age range and could potentially under or overestimate their true skills. Similarly, the home stimulation measure only covers six very basic caregiver-child interactions over a short time (Lu et al., 2020). Fourth, given that we only possess a binary variable indicating whether children have literacy/numeracy skills, we are unable to examine the distribution across different quantile levels. This limitation restricts our ability to analyze the full range of performance within each gender. To illustrate, previous research by Autor et al. (2023) and Bossavie & Kanninen (2018) has demonstrated that males often exhibit a wider dispersal of education outcomes compared to females. Consequently, other research has found that

the overall underachievement of males can be primarily attributed to a subset of poorly-performing male students who significantly lower the average scores.

7. Conclusion

Gender educational disparities have far-reaching implications for lifelong academic achievement, productivity, and well-being. Given the pivotal role of education, it is crucial to understand potential disparities at pre-school ages. Education disparities could pose significant obstacles to human capital development, particularly for marginalized groups facing violence, poor health outcomes, and economic instability (Welmond & Gregory, 2021). Since the early development of numeracy and literacy skills in children serves as a critical predictor of their subsequent learning and academic performance, our study aimed to examine gender differences among children in their early educational skills and understand disparities at an early age in developing countries.

To establish evidence of gender disparities in early childhood educational outcomes, we analyze data from national household surveys conducted between 2011 and 2019 across 25 Sub-Saharan African countries. Our findings indicate that girls have a slightly higher probability of outperforming boys in at least two out of three learning skills, with an estimated difference of approximately two percentage points. While there is variation across countries, girls outperform boys in most of the countries in our data. Our analysis further reveals that maternal literacy significantly contributes to the gender gap.

One of the key contributions of this paper is the use of rich individual-level data to document the gender gap at the child level. While the paper does not attempt to explain the large cross-country differences in gender gaps, it does highlight the variation by country and presents the gender gap separately for countries with higher versus lower levels of early childhood education program attendance. However, conducting further analysis at the country level would allow for a more detailed examination of gender differences. While this study emphasizes early educational underachievement among boys, it does not imply that girls' educational access and outcomes are without challenges. In specific contexts, girls may lag behind boys in early literacy and numeracy skills, posing ongoing concerns, especially in economically disadvantaged countries. Moreover, our estimates reveal that only a mere 20 percent of three- and four-year-old children in our sample possessed these early educational skills, highlighting the need to improve early learning

opportunities for both genders. We also acknowledge the struggles that all children – boys and girls alike – face in the education system within sub-Sahara Africa and more broadly across other non-education domains.

We acknowledge that our analysis is based on cross-sectional comparisons, and therefore, it would be valuable for future research to incorporate a longitudinal perspective. A time series of the gender gap would enhance our understanding of its evolution and shed light on whether increased educational opportunities for both boys and girls accompany it. This information is critical for comprehending the dynamics of the gender gap, including whether girls are progressing faster than boys or if boys are experiencing a decline in performance over time.

References

Akresh, R., Halim, D., & Kleemans, M. (2023). Long-term and intergenerational effects of education: Evidence from school construction in Indonesia. *The Economic Journal*, *133*(650), 582–612.

Anders, Y., Grosse, C., Rossbach, H.-G., Ebert, S., & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement*, *24*(2), 195–211.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics*, *11*(3), 338–381.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2023). Males at the tails: How socioeconomic status shapes the gender gap. *The Economic Journal*, *133*(656), 3136–3152.

Autor, D., & Wasserman, M. (2013). Wayward sons: The emerging gender gap in labor markets and education. *Third Way Report*, *20013*.

Bago, J.-L., Souratié, W. M., Ouédraogo, E., Lompo, M. L., Ouédraogo, M., & Perrault, N. (2019). *Is there an early gender gap in Ghanaian children development? Evidence from 3-4 years old boys and girls* (MPRA Paper 95876). University Library of Munich.

Bando, R., Lopez-Boo, F., Fernald, L., Gertler, P., & Reynolds, S. (2024). Gender Differences in Early Child Development: Evidence from Large-Scale Studies of Very Young Children in Nine Countries. *Journal of Economics, Race, and Policy*. https://doi.org/10.1007/s41996-023-00131-1

Bertocchi, G., & Bozzano, M. (2020). *Gender gaps in education*. Springer.

Bharadwaj, P., De Giorgi, G., Hansen, D., & Neilson, C. A. (2016). The Gender Gap in Mathematics: Evidence from Chile. *Economic Development and Cultural Change*, *65*(1), 141–166. https://doi.org/10.1086/687983

Björkman-Nyqvist, M. (2013). Income shocks and gender gaps in education: Evidence from Uganda. *Journal of Development Economics*, *105*, 237–253. https://doi.org/10.1016/j.jdeveco.2013.07.013

Bossavie, L., & Kanninen, O. (2018). *What explains the gender gap reversal in education? The role of the tail hypothesis*. The World Bank.

Brenøe, A. A., & Lundberg, S. (2018). Gender gaps in the effects of childhood family environment: Do they persist into adulthood? *European Economic Review*, *109*, 42–62.

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annu. Rev. Sociol*, *34*, 319–337.

Buhl-Wiggers, J., Jones, S., & Thornton, R. (2021). Boys lagging behind: Unpacking gender differences in academic achievement across East Africa. *International Journal of Educational Development*, *83*, 102382.

Chuan, A., List, J. A., Samek, A., & Samujjwala, S. (2022). Parental Investments in Early Childhood and the Gender Gap in Math and Literacy. *AEA Papers and Proceedings*, *112*, 603–608.

Cobb-Clark, D. A., & Moschion, J. (2017). Gender gaps in early educational achievement. *Journal of Population Economics*, *30*, 1093–1134.

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

Dickerson, A., McIntosh, S., & Valente, C. (2015). Do the maths: An analysis of the gender gap in mathematics in Africa. *Economics of Education Review*, *46*, 1–22. https://doi.org/10.1016/j.econedurev.2015.02.005

Diniz, E., Brandão, T., Monteiro, L., & Veríssimo, M. (2021). Father Involvement During Early Childhood: A Systematic Review of the Literature. *Journal of Family Theory & Review*, *13*(1), 77–99. https://doi.org/10.1111/jftr.12410

DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, *41*(1), 1–15.

Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *Sociological quarterly*, *46*(2), 299–321.

García, J. L., Heckman, J. J., Leaf, D. E., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, *128*(7), 2502–2541.

García, J. L., Heckman, J. J., & Ronda, V. (2021). *The lasting effects of early childhood education on promoting the skills and social mobility of disadvantaged african americans*. National Bureau of Economic Research.

Gerrand, P., & Nduna, M. (2021). Father Absence in the Lives of Female African Youth Living in Mpumalanga, South Africa: Christianity a Coping Strategy that Builds and Strengthens Resilience. *Social Work & Christianity*, *48*(2), 183. https://doi.org/10.34043/SWC.V48I2.189

Gilliam, W. S. & others. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. Foundation for Child Development New York, NY.

Glick, P., & Sahn, D. E. (2000). Schooling of girls and boys in a West African country: The effects of parental education, income, and household structure. *Economics of Education Review*, *19*(1), 63–87. https://doi.org/10.1016/S0272-7757(99)00029-1

Grant, M. J., & Behrman, J. R. (2010). Gender Gaps in Educational Attainment in Less Developed Countries. *Population and Development Review*, *36*(1), 71–89.

Hancioglu, A., & Arnold, F. (2013). Measuring coverage in MNCH: tracking progress in health for women and children using DHS and MICS household surveys. *PLoS medicine*, *10*(5), e1001391.

Heath, R., & Mobarak, A. M. (2015). Manufacturing growth and the lives of Bangladeshi women. *Journal of development Economics*, *115*, 1–15.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052–2086.

Henry, J. B., Julion, W. A., Bounds, D. T., & Sumo, J. (2020). Fatherhood Matters: An Integrative Review of Fatherhood Intervention Research. *The Journal of School Nursing*, *36*(1), 19–32. https://doi.org/10.1177/1059840519873380

Jeong, J., Franchett, E. E., Ramos de Oliveira, C. V., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS medicine*, *18*(5), e1003602.

Kahn, S., & Ginther, D. (2017). *Women and STEM*. National Bureau of Economic Research.

Lai, F. (2010). Are boys left behind? The evolution of the gender achievement gap in Beijing's middle schools. *Economics of Education Review*, *29*(3), 383–399. https://doi.org/10.1016/j.econedurev.2009.07.009

Lloyd, C. B., & Blanc, A. K. (1996). Children's Schooling in sub-Saharan Africa: The Role of Fathers, Mothers, and Others. *Population and Development Review*, *22*(2), 265–298. https://doi.org/10.2307/2137435

Loizillon, A., Petrowski, N., Britto, P., & Cappa, C. (2017). *Development of the Early Childhood Development Index in MICS surveys* (Data and Analytics Section, Division of Data, Research and Policy 6). UNICEF.

Lu, C., Cuartas, J., Fink, G., McCoy, D., Liu, K., Li, Z., Daelmans, B., & Richter, L. (2020). Inequalities in early childhood care and development in low/middle-income countries: 2010–2018. *BMJ Global Health*, *5*(2), e002314.

Lundberg, S. (2022). *Father Absence and the Educational Gender Gap* (RePEc 10814). UC Santa Barbara. https://escholarship.org/uc/item/1nw6459h

Lundberg, S. J. (2017). *Father absence and the educational gender gap*.

Madhavan, S., Myroniuk, T. W., Kuhn, R., & Collinson, M. A. (2017). Household structure vs. composition: Understanding gendered effects on educational progress in rural South Africa. *Demographic research*, *37*, 1891–1916. https://doi.org/10.4054/DemRes.2017.37.59

Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H., & Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. *Early Childhood Research Quarterly*, *36*, 521–536. https://doi.org/10.1016/j.ecresq.2015.12.021

McCoy, D. C., Peet, E. D., Ezzati, M., Danaei, G., Black, M. M., Sudfeld, C. R., Fawzi, W., & Fink, G. (2016). Early childhood developmental status in low-and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLoS medicine*, *13*(6), e1002034.

McCoy, D. C., Salhi, C., Yoshikawa, H., Black, M., Britto, P., & Fink, G. (2018). Home-and center-based learning opportunities for preschoolers in low-and middle-income countries. *Children and Youth Services Review*, *88*, 44–56.

Mullis, I. V., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading.* ERIC.

Nakajima, N., Jung, H., Pradhan, M. P., Hasan, A., Kinnell, A., & Brinkman, S. (2016). Gender gaps in cognitive and non-cognitive skills in early primary grades: Evidence from rural Indonesia. *World Bank Policy Research Working Paper*, *7833*.

Ng'ang'a, A., Mureithi, L. P., & Wambugu, A. (2018). Mathematics gender gaps in Kenya: Are resource differentials between boys and girls to blame? *Cogent Education*, *5*(1), 1564163. https://doi.org/10.1080/2331186X.2018.1564163

Owens, J. (2016). Early childhood behavior problems and the gender gap in educational attainment in the United States. *Sociology of education*, *89*(3), 236–258.

Reeves, R. V. (2022). *Of Boys and Men: Why the Modern Male Is Struggling, Why It Matters, and What to Do about It*. Brookings Institution Press.

Rühle, R. (2022). *Mind the Gap: An Analysis of Gender Differences in Mathematics and Science Achievement in South Africa*. Department of Economics, University of Stellenbosch.

Saito, M. (2011). Trends in the magnitude and direction of gender differences in learning outcomes. *SACMEQ: UNESCO International Institute for Educational Planning*.

Soto-Calvo, E., & Sánchez-Barrioluengo, M. (2016). *The influence of early literacy competences on later mathematical attainment: Evidence from TIMSS & PIRLS 2011*.

Spaull, N., & Makaluza, N. (2019). Girls do better: The pro-female gender gap in learning outcomes in South Africa 1995–2018. *Agenda*, *33*(4), 11–28.

Suryadarma, D. (2015). Gender differences in numeracy in Indonesia: Evidence from a longitudinal dataset. *Education Economics*, *23*(2), 180–198. https://doi.org/10.1080/09645292.2013.819415

Timaeus, I. M., & Boler, T. (2007). Father figures: The progress at school of orphans in South Africa. *AIDS*, *21*, S83. https://doi.org/10.1097/01.aids.0000300539.35720.a0

UNESCO. (2015). *Education for all 2000–2015: Achievements and challenges*. UNESCO.

Wasserman, M. (2020). The Disparate Effects of Family Structure. *The Future of Children*, *30*(1), 55–82.

Welmond, M. J., & Gregory, L. (2021). *Educational underachievement among boys and men*.

Wilder, G. Z., & Powell, K. (1989). *Sex differences in test performance: A survey of literature* (Bd. 89). Citeseer.

World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise* [Text/HTML]. http://www.worldbank.org/en/publication/wdr2018

Zuze, L., Reddy, B. V., Visser, M., Winnaar, L., & Govender, A. (2017). *TIMSS 2015 Grade 9 national report: Understanding mathematics and science achievement amongst Grade 9 learners in South Africa*. HSRC Press.

Zuze, T. L., & Reddy, V. (2014). School resources and the gender reading literacy gap in South African schools. *International Journal of Educational Development*, *36*, 100–107. https://doi.org/10.1016/j.ijedudev.2013.10.002

Appendix

A1 Table. List of countries included in the analysis and the year of the survey

N	Countries	DHS 6	DHS 7	MICS5	MICS6	Sample Size
1	Benin	2017		2014		11,186
2	Burundi	2016				8,065
3	Cameroon		2011	2014		5,379
4	Central African Republic				2018-2019	2,963
5	Chad	2014			2019	12,005
6	Congo			2014-2015		3,145
7	Coted Ivoire			2016		3,207
8	DR Congo	2013			2017-2018	12,809
9	Gambia				2018	3,740
10	Ghana				2017-2018	3,121
11	Guinea			2016		2,562
12	Guinea Bissau			2014	2018-2019	2,892
13	Lesotho				2018	515
14	Madagascar				2018	4,389
15	Malawi			2013-2014	2019-2020	11,868
16	Mali			2015		5,419
17	Mauritania			2015		3,406
18	Nigeria			2016-2017		9,970
19	Rwanda	2014				4,006
20	SaoTome			2014	2019	1,372
21	Senegal	2017		2015-2016		8,449
22	Swaziland			2014		774
23	Togo				2017	1,713
24	Uganda	2016		2014		8,236
25	Zimbabwe				2019	5,340

A2 Table. Definition of Variables

Variable Name	Definition
Panel A. Children Variables	
Female	1 = Female
Age	Child's age in complete years
Identify 10 letters	1 = children can identify or name at least 10 letters of the alphabet
Read 4 words	1 = children can read at least four simple, common words
Identify 10 numbers	1 = children knows the name and recognize the symbols for all numbers from 1 to 10
Have 3 skills	1 = children has the three literacy/numeracy skills
Early Childhood Education Program	1 = child attends an early childhood education program
Panel B.Parental Home Stimulation Activities	
Read to children	1 = at least one parent read books aloud to their child in the past 3 days.
Told stories to children	1 = at least one parent told stories to the child in the past 3 days
Sang songs to children	1 = at least one parent sang songs (including lullabies) to the child in the past 3 days
Took child outside home	1 = at least one parent took the child outside the home, compound, yard or enclosure in the past 3 days
Play with child	1 = at least one parent played with the child in the past 3 days.
Named/counted/drew with child	1 = at least one parent named, counted, or drew things with child in the past 3 days
N of Parental home stimulation activities	Number of home stimulation activities carried out by at least one parent
Panel C. Parent and Household Characteristics	
Mother's Age	Women's age in complete years
Literate mother	1 = Able to read whole sentence
Number of HH members	Number of household members
Father is not part of the household	1 = father is absent from the household
Poor	1 = wealth index quintile is equal to first and second lowest quintile, out of 5
Urban	1 = Urban, 0 = Rural

A3 Table. Gender Effects on Preschool Literacy/Numeracy by Mother's Literacy

	Identify 10 letters	Read 4 words	Identify 10 numbers
	(1)	(2)	(3)
	Panel A. Literate Mother Sample		
Female	0.031***	0.009	0.028***
	(0.007)	(0.005)	(0.007)
Overall mean	0.28	0.18	0.30
Observations	50,071	50,071	50,071
R-squared	0.185	0.120	0.173
	Panel B. Illiterate Mother Sample		
Female	0.008	0.003	0.010**
	(0.005)	(0.003)	(0.004)
Overall mean	0.10	0.09	0.11
Observations	86,460	86,460	86,460
R-squared	0.088	0.043	0.070

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. Panel A consists of observations where the mother can read a full sentence, while Panel B includes a sample of mothers who cannot read or can only read parts of a sentence. The independent variables in Columns 1-3 are dummy variables that measure whether the child shows early literacy or numeracy skills based on their mother's response. All columns include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's absence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A4 Table. Gender Effects on Preschool Literacy/Numeracy – Village FEs

	Identify 10 letters	Read 4 words	Identify 10 numbers
	(1)	(2)	(3)
Female	0.014***	0.005**	0.014***
	(0.003)	(0.003)	(0.003)
Overall mean	0.17	0.12	0.18
Observations	135,413	135,413	135,413
R-squared	0.416	0.346	0.380
Country FE	YES	YES	YES
Survey-Year FE	YES	YES	YES
Village FE	YES	YES	YES
Controls	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In columns 1, 2, and 3 the dependent variables are dummy variables that measure whether the child shows early literacy or numeracy skills based on their mother's response. All columns include the following controls: child's age, mother's age, mother's education level (at least secondary education), household size, father's presence in the household, household poverty status, and residence area (urban/rural). All regressions include population weights, country, survey-year, and village fixed effects. Standard errors clustered at the village level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A5 Table. Heterogeneity of the Gender Gap – Identify 10 letters

	Identify 10 letters			
	(1)	(2)	(3)	(4)
Female	0.008	0.017***	0.023**	0.016**
	(0.005)	(0.005)	(0.008)	(0.008)
Literate mother	0.120***	0.132***	0.132***	0.132***
	(0.024)	(0.025)	(0.025)	(0.025)
Father's absence in the household	-0.001	-0.002	-0.001	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)
Poor	-0.081***	-0.081***	-0.074***	-0.081***
	(0.018)	(0.018)	(0.016)	(0.018)
Urban	0.082***	0.082***	0.082***	0.082***
	(0.011)	(0.011)	(0.011)	(0.013)
Mother's literacy x female	0.024***			
	(0.005)			
Father's absence in the household x female		0.001		
		(0.008)		
Poor x Female			-0.013**	
			(0.006)	
Urban x Female				0.001
				(0.007)
Observations	136,531	136,531	136,531	136,531
R-squared	0.158	0.157	0.157	0.157
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In addition to the explanatory variables and its interaction term with child's gender, all columns include the following controls: child's age, mother's age, and household size. All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A6 Table. Heterogeneity of the Gender Gap – Read four words

	Read 4 Words			
	(1)	(2)	(3)	(4)
Female	0.003	0.004	0.006	0.007
	(0.003)	(0.003)	(0.005)	(0.004)
Literate mother	0.083***	0.087***	0.087***	0.087***
	(0.021)	(0.021)	(0.021)	(0.021)
Father's absence in the household	-0.004	-0.007	-0.004	-0.004
	(0.004)	(0.005)	(0.004)	(0.004)
Poor	-0.055***	-0.055***	-0.055***	-0.055***
	(0.013)	(0.013)	(0.013)	(0.013)
Urban	0.043***	0.043***	0.043***	0.045***
	(0.009)	(0.009)	(0.009)	(0.009)
Mother's literacy x female	0.006			
	(0.005)			
Father's absence in the household x female		0.005		
		(0.005)		
Poor x Female			-0.001	
			(0.006)	
Urban x Female				-0.004
				(0.006)
Observations	136,531	136,531	136,531	136,531
R-squared	0.081	0.081	0.081	0.081
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In addition to the explanatory variables and its interaction term with child's gender, all columns include the following controls: child's age, mother's age, and household size. All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A7 Table. Heterogeneity of the Gender Gap – Identify 10 numbers

	Identify 10 Numbers			
	(1)	(2)	(3)	(4)
Female	0.010**	0.016***	0.022***	0.013**
	(0.004)	(0.004)	(0.006)	(0.005)
Literate mother	0.125***	0.134***	0.134***	0.134***
	(0.025)	(0.026)	(0.026)	(0.026)
Father's absence in the household	-0.002	-0.004	-0.002	-0.002
	(0.004)	(0.005)	(0.004)	(0.004)
Poor	-0.084***	-0.085***	-0.079***	-0.085***
	(0.017)	(0.017)	(0.015)	(0.017)
Urban	0.083***	0.083***	0.083***	0.076***
	(0.010)	(0.010)	(0.010)	(0.009)
Mother's literacy x female	0.018**			
	(0.007)			
Father's absence in the household x female		0.005		
		(0.006)		
Poor x Female			-0.012**	
			(0.005)	
Urban x Female				0.014*
				(0.007)
Observations	136,531	136,531	136,531	136,531
R-squared	0.155	0.155	0.155	0.155
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

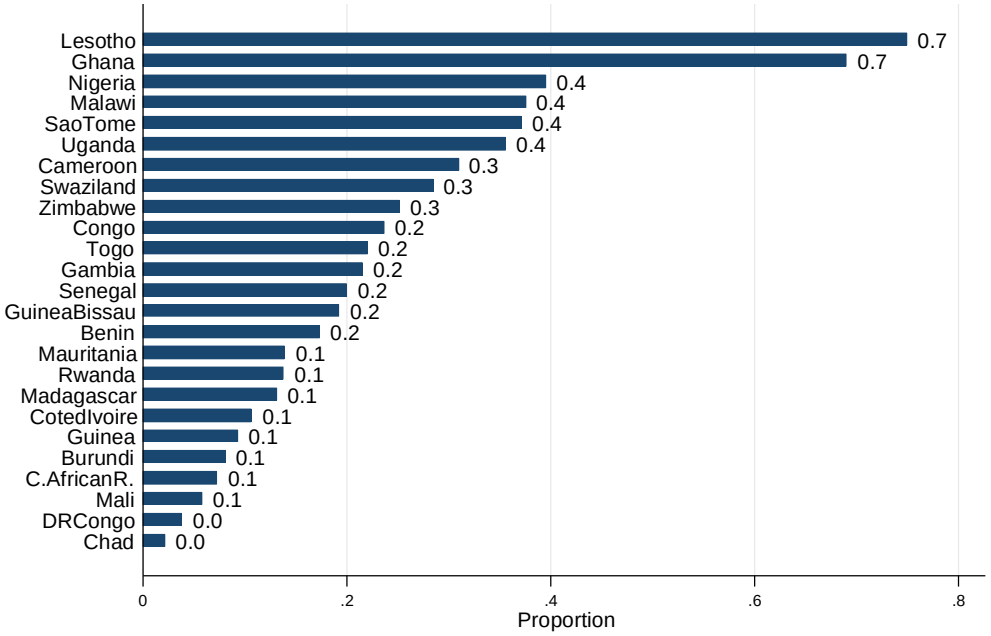
Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In addition to the explanatory variables and its interaction term with child's gender, all columns include the following controls: child's age, mother's age, and household size. All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A8 Table. Heterogeneity on Early Childhood Education and Home Stimulation

	Early Childhood Education Program					At least one home stimulation activity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.006	0.019**	0.010	0.013*	0.012	0.002	0.011	0.006	0.005	0.005
	(0.005)	(0.009)	(0.007)	(0.007)	(0.012)	(0.005)	(0.009)	(0.005)	(0.004)	(0.011)
Literate mother	0.153***	0.160***	0.160***	0.160***	0.152***	0.081***	0.086***	0.086***	0.086***	0.081***
	(0.028)	(0.029)	(0.029)	(0.029)	(0.028)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)
Father's absence in HH	0.009	0.004	0.009	0.009	0.005	-0.077***	-0.080***	-0.077***	-0.077***	-0.079***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.013)	(0.014)	(0.013)	(0.013)	(0.014)
Poor	-0.114***	-0.114***	-0.116***	-0.114***	-0.116***	-0.059***	-0.059***	-0.059***	-0.059***	-0.060***
	(0.021)	(0.021)	(0.019)	(0.021)	(0.019)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)
Urban	0.095***	0.095***	0.095***	0.098***	0.099***	0.047***	0.047***	0.047***	0.046***	0.047***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Mother's literacy x female	0.014**				0.016***	0.010				0.010
	(0.006)				(0.006)	(0.011)				(0.011)
Father's absence in HH x female		0.010*			0.009		0.006			0.005
		(0.006)			(0.006)		(0.010)			(0.010)
Poor x Female			0.003		0.005			-0.001		0.002
			(0.006)		(0.007)			(0.006)		(0.007)
Urban x Female				-0.005	-0.007				0.001	-0.000
				(0.007)	(0.008)				(0.008)	(0.009)
Observations	136,531	136,531	136,531	136,531	136,531	136,531	136,531	136,531	136,531	136,531
R-squared	0.261	0.261	0.261	0.261	0.261	0.058	0.058	0.058	0.058	0.058
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: The sample consists of surviving and co-resident children aged 3 and 4 from DHS and MICS respondents. In addition to the explanatory variables and its interaction term with child's gender, all columns include the following controls: child's age, mother's age, and household size. All regressions include population weights, country, and survey-year fixed effects. Standard errors clustered at the country level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A1 Figure: Percent of 3- and 4-Year-Old Children attending Early Childhood Education



Notes: This figure shows the percentage of children from the data who attend an early childhood education program.