WPS7833

Policy Research Working Paper 7833

Gender Gaps in Cognitive and Non-Cognitive Skills in Early Primary Grades

Evidence from Rural Indonesia

Nozomi Nakajima
Haeil Jung
Menno Pradhan
Amer Hasan
Angela Kinnell
Sally Brinkman

WORLD BANK GROUP
Education Global Practice Group
September 2016

Policy Research Working Paper 7833

Abstract

This paper examines gender gaps in cognitive and non-cognitive skills among a sample of more than 10,000 children between the ages of 6 and 9 in rural Indonesia. In terms of cognitive skills, the analysis finds evidence of gender gaps favoring girls at each age in test scores of language (0.158–0.252 standard deviations) and mathematics (0.155–0.243 standard deviations) in the early years of primary school. Girls also perform significantly better than boys in non-cognitive skills, with higher scores on the social competence (0.086–0.247 standard deviations) and emotional maturity domains (0.213–0.296 standard deviations) of the Early Development Instrument, a finding consistent with research from high-income countries. Decomposition analyses are used to investigate the extent to which enrollment patterns in preschool and primary school as well as parenting practices contribute to these gender gaps in cognitive and non-cognitive skills. Standard decomposition approaches are extended to correct for selection on observables. The findings show that gender differences in enrollment patterns play a role in explaining gender gaps in test scores, while differences in parenting practices do not. However, the relative contribution of observed factors to gender gaps depends on the available quality of preschool services in the child's village and whether the outcome of interest is cognitive or non-cognitive skills.

This paper is a product of the Education Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at ahasan1@worldbank.org.

Gender Gaps in Cognitive and Non-Cognitive Skills in Early Primary Grades: Evidence from Rural Indonesia

Nozomi Nakajima[1], Haeil Jung[2], Menno Pradhan[3], Amer Hasan[1], Angela Kinnell[4], and Sally Brinkman[5] *

JEL classification codes: I20, I24

Keywords: gender gaps, primary school, Early Childhood Education and Development (ECED), Indonesia, cognitive and non-cognitive skills

[1] Education Global Practice, World Bank, Washington, D.C., USA
[2] Department of Public Administration, Korea University, Seoul, South Korea
[3] University of Amsterdam and VU University, Amsterdam, the Netherlands
[4] School of Psychology, University of Adelaide, Adelaide, Australia
[5] Fraser Mustard Centre, Telethon Kids Institute, University of Western Australia, Perth, Australia
* Acknowledgements: We would like to thank Dedy Junaedi, Upik Sabainingrum, Anas Sutisna, Lulus Kusbudiharjo and Mulyana for managing the fieldwork. Mayla Safuro Lestari Putri and Mulyana provided excellent research assistance. This collaboration has been supported by the DIME-administered IE2I TF (TF018488). Data collection was partially funded by the Government of the Kingdom of the Netherlands through the Dutch Education Support Program Trust Fund (TF057272).

1. Introduction

Investigating differences in human capital accumulation between males and females has been the focus of much research over the years. Gender featured in two of the eight Millennium Development Goals: to achieve universal primary education for all boys and girls (Goal 2) and to promote gender equality and empower women, with a focus on ending gender disparity in primary and secondary education (Goal 3). These goals were to be achieved by 2015, but gender disparities in these areas still exist in many countries around the world. Perhaps unsurprisingly, gender equality continues to feature prominently in the Sustainable Development Goals. Gender equality—particularly in education—can help to reduce child mortality, improve health outcomes, improve the productivity of communities and countries, and help them to improve their financial position and minimize poverty (OECD 2015, World Bank 2011).

Research has shown that gender differences in educational achievement emerge in the early years of school and can persist into adulthood (Anderson 2008). There is also a growing body of evidence suggesting that non-cognitive skills observed in early childhood affect academic performance and labor market outcomes in later years (Cunha and Heckman 2008, Cunha et al. 2010). As a result, there is considerable interest in understanding the extent to which gender gaps exist in cognitive and non-cognitive skills in the early years, and what factors may explain these observed gender gaps. However, research from developing countries on this topic is still relatively sparse.

In this paper, we focus on rural Indonesia and investigate gender differences across an array of developmental domains for children in the first few grades of primary school. We also contribute to the literature on gender gaps by examining whether early schooling and parenting practices account for this observed gender gap and, if so, to what extent. Our rich data set allows us to measure not only the quantity of early schooling but also its quality. In addition, the data set also contains information on the parenting practices children experience. Furthermore, this paper implements an extension of the Oaxaca-Blinder decomposition method that takes into account moderating factors to explore the extent to which early schooling and parenting practices contribute to the gender gaps we observe in the first few years of primary school.

Our analysis is based on data collected as part of an impact evaluation of an intervention, which expanded access to preschool services in parts of rural Indonesia. As part of this study we collected measures of cognitive and non-cognitive skills of children in the early grades of primary

school along with full enrollment histories in preschool and primary school. We also collected information on parenting practices, the quality of preschools in the village, and other child and family background characteristics.

We find substantial gender gaps in both cognitive and non-cognitive skills in early primary grades, with substantial female advantage in language and mathematics assessments as well as in emotional maturity and social competence. We also find that the rate and duration of enrollment in preschools for girls is higher than for boys in the prime preschool-going ages of 3 to 5. In addition, gender gaps are observed in parenting practices for children age 7 to 9, suggesting that there are some gender differences in parent-child interactions in rural Indonesia. In our preferred decomposition analyses, we find that these observed gender gaps in cognitive and non-cognitive skills are in part explained by one's early schooling environment.

The rest of the paper is organized as follows. In the next section, we discuss relevant literature and in section 3, we describe the country context. The data and key variables are described in section 4. In section 5, we describe the gender differences in cognitive and non-cognitive skills in the early years of primary school and explore whether there are differences in early schooling and family environment between girls and boys. In section 6, we use decomposition analyses to investigate the extent to which these school and family factors contribute to the gender gaps we observe in the first few years of primary school. The paper concludes by discussing the results in relation to prior studies on the topic of gender gaps in early childhood developmental outcomes.

2. Literature review

Early achievement gaps can lead to large educational disparities in later life (Currie and Thomas 1999, Paxson and Schady 2005). As a result, there is considerable interest in understanding when these gaps in child development appear. Recent studies from developed countries have examined the emergence of gender gaps in cognitive and non-cognitive outcomes in the first year of schooling (Anderson 2008, Cornwell et al. 2013, DiPrete and Jennings 2012, Fryer and Levitt 2010, Janus and Duku 2007). Evidence from nationally representative data of kindergarten students in the United States shows that while there are no mean differences in math scores between boys and girls in kindergarten, girls lose ground to boys by more than 0.2 standard deviation (SD) by fifth grade (Fryer and Levitt 2010). Other studies using the same data set also

show that girls significantly outperform boys on reading tests—by 0.162 SD in kindergarten—and that this female advantage persists into the fifth grade (Cornwell et al. 2013).

Recent evidence from developing countries also shows gender gaps in educational achievement. Dickerson et al. (2015) find that on average, boys have a 0.1 SD advantage over girls in mathematics in fifth grade in 19 African counties but the magnitude of the gap varies widely from a non-significant 0.06 SD in Mauritius to as much as a 0.34 SD in Tanzania. Similarly, boys in Chile score 0.08 SD higher in mathematics than girls in fourth grade, and this gap increases to 0.2 SD by grade eight (Bhardwaj et al. 2015). In Indonesia, the gender gap in maths scores in primary school is in the opposite direction, with girls scoring 0.08 SD higher than boys at age 11 (Suryadarama 2015). Moreover, analysis using international assessment data suggests significant gender gaps in secondary school among low- and middle-income countries, with boys scoring 0.25 SD higher than girls in mathematics at age 15 (Bhardwaj 2015). A question that emerges from the existing literature from developing counties is whether gender gaps in cognitive skills emerge as early as those observed in high income countries (i.e., the first year of schooling).[6]

Researchers have also paid increasing attention to gender gaps in non-cognitive skills given the growing evidence suggesting that non-cognitive skills observed in early childhood affect academic performance and labor market outcomes in later years (Cunha and Heckman 2008, Cunha et al. 2010). In the United States, girls consistently score higher than boys from kindergarten to grade 5 on a teacher-reported measure of children's socio-emotional development (Cornwell et al. 2013, DiPrete and Jennings 2012). Similarly, results from Australia and Canada show that in kindergarten, girls outperform boys on the Early Development Instrument (EDI)—a holistic measure of child development that includes measures of non-cognitive skills. Gender gaps in the EDI are particularly pronounced in the social competence domain (i.e., children's ability to cooperate with others and follow rules) and the emotional maturity domain (i.e., children's ability to deal with feelings at the age-appropriate level) (Australian Government 2013, Janus and Duku 2007). To our knowledge, research from developing countries has yet to examine the early emergence of gender gaps in non-cognitive skills.

[6] A cross-country study from Ethiopia, India, Peru and Vietnam suggests that there were no gender gaps in general cognitive skills at age 4 to 6 (Cueto et al. 2009) but gender difference grew significant during middle childhood (age 12), with male advantage in cognitive achievement in Ethiopia, India and Peru, and female advantage in Vietnam (Dercon and Singh 2011).

What explains these observed differences in early outcomes between girls and boys? A number of recent studies focus on the role of schooling and family environments in explaining observed gender gaps in cognitive and non-cognitive skills (Autor et al. 2016).[7] In the case of gender gaps observed in the first few years of primary school, early schooling experiences are likely to play a role. Research on the effect of preschool duration suggests that children with longer exposure to preschool demonstrate advantages over children with shorter exposure (Arteaga et al. 2014, Domitrovich et al. 2013, Loeb et al. 2007). A review of 30 preschool interventions in 23 countries shows that early childhood education programs lasting one to three years had an average effect size of 0.312 SD, compared to only 0.196 SD for programs lasting less than one year (Nores and Barnett 2010). In addition to duration, the literature points to the importance of quality of early childhood education programs in sustaining impacts on children's cognitive and non-cognitive skills (Engle et al. 2011). Thus, if girls and boys are exposed to different quantity and quality of preschools, we would expect to see these early schooling factors explain part of the gender gaps observed in the early years of primary school.

In addition to early schooling factors, children's family environment is likely to play a role in the emerging gender gap in cognitive and non-cognitive skills. Boys may react differently than girls to the family environment, and/or parents may adjust their parenting practices depending on the gender of the child. Using panel data from the United States, Owens (2013) shows that the higher levels of early childhood behavior problems exhibited by boys—which account for up to 25 percent of the gender gap in educational attainment at age 22—are largely explained by gender differences in parent-child interactions. Specifically, boys respond more negatively than girls when they are exposed to parental conflict and parental harshness. Baker and Milligan (2013) also provide evidence from the U.S., Canada and the U.K. that parents spend more time with girls than boys in parental teaching activities such as reading and the use of numbers and letters, and these higher parental inputs for girls account for 23 percent of the gender gap in reading abilities in preschool. In developing countries, differences in parental expectations toward girls and boys are widely documented (see Bhardwaj et al. 2014 for a comprehensive review) and as such, gender

[7] In addition to school and family factors, a range of other explanatory factors has been explored in the gender gaps literature. For example, Goldin et al. (2006) show that macro-level social and economic changes are a key explanation for the gender gap in educational attainment, while Bertrand (2011) provides an overview of how psychological and socio-psychological factors explain gender differences in educational and labor market outcomes.

differences in the quality of parent-child interactions during early childhood are likely to explain part of the observed gender gaps in children's cognitive and non-cognitive skills in the early years.

In light of the literature reviewed above, in this paper, we investigate gender differences across an array of cognitive and non-cognitive skills for children in the first few grades of primary school, and examine the extent to which early schooling experiences and parenting practices account for this observed gender gap in rural Indonesia.

3. Country context

Indonesia has the fourth largest education system in the world with over 50 million students, 2.6 million teachers, and more than 250,000 schools. In 2012, the government expenditure on education was 4% of its GDP. Net enrollment rates in primary, secondary, and tertiary education are 92%, 77% and 32% respectively (World Bank 2015). There are no significant differences in primary and secondary education enrollment rates between girls and boys (Suryadarma 2015).

However, results of educational achievement data during primary and secondary schooling show some evidence of gender gaps. Girls significantly outperform boys in reading in the fourth grade (Mullis et al. 2012) and by age 15, this female advantage is equivalent to approximately 10 additional months of schooling (OECD 2015). In contrast, results in mathematics are mixed. Using a longitudinal household survey from Indonesia, Suryadarma (2015) finds that girls score 0.08 SD higher in numeracy tests than boys at age 11 and this gap increases to 0.19 SD when the sample of children are 18 years-old in 2007. In contrast, results from the most recent round of PISA show that the difference between boys and girls in mathematics at around age 15 is small in magnitude and not statistically significant (OECD 2015). Thus, the existing evidence of educational outcomes of Indonesian children shows mixed evidence of gender gaps during late primary and secondary school. This paper adds to the gender gap literature in Indonesia by focusing on the early childhood years and examining both cognitive and non-cognitive outcomes.

4. Data and measures

4.1 Data

Data for this study were collected in 2013 as part of an impact evaluation of the Indonesia ECED Project (see Pradhan et al., 2013 for a detailed description of the study protocol). The Indonesia ECED Project was designed to improve poor children's school readiness by establishing community-based early childhood education programs.

4.2 Measures

In this section we describe the main dimensions along which we investigate gender gaps in cognitive and non-cognitive outcomes. We also describe the variables used to measure school and family environments, which are our two key predictors of the observed gender gaps in early childhood. Summary statistics of these key variables are all shown in Table 1.

Test scores

Our sample consists of 10,858 primary school students between six and nine years of age in 2013. At the time these children were enrolled in grades 1 through 4 of primary school. We administered a test of Bahasa Indonesia, mathematics and abstract reasoning to these children. The test items for Bahasa and mathematics are from a battery of questions that align with the national curriculum for lower primary school grades and the test items for abstract reasoning are based on the Raven's colored Progressive Matrices. Two versions of the test were administered: an easier test for 6 and 7 year-olds and a more difficult test for 8 and 9 year-olds. There were 39 common items across the two versions of the test, which we use in our analysis.[8] Table 1 shows the number of items each child answered correctly. On average girls score slightly higher than boys on the test overall with differences in the language and mathematics sections but no difference in the abstract reasoning section.

For our analysis, we standardize the test scores using the mean and standard deviation of children who were age 6 since the standard deviation of the raw test scores are quite similar across the four ages.

Early Development Instrument

Child development outcomes in early childhood are measured using the Early Development Instrument (EDI), which has been demonstrated as a valid and reliable measure of child development (Janus and Offord 2007, Forget-Dubois et al. 2007). The EDI was adapted and translated for use in the Indonesia ECED Project by the authors and members of the research team (Brinkman et al. 2015). There are five domains in the EDI: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills

[8] An analysis of the test items using item response theory (IRT) shows similar levels of item difficulty for boys and girls. See Nakajima and Hasan (2016) for details.

and general knowledge.[9] Each domain is scored from 1 (low) to 10 (high). For the purpose of measuring gender gaps in non-cognitive skills, we are particularly interested in the social competence and emotional maturity domains of the EDI. Data are available for 8,653 children who were age 8 and below.[10]

Summary statistics for the raw EDI score in 2013 are shown in Table 1. On average, the EDI scores of these children (ages 6 to 8) show a slight female advantage with girls scoring higher than boys in all five domains.

Enrollment

Information on enrollment history in preschool and primary school for each academic year between 2008 and 2013 was collected from the mother or main caregiver of the 10,858 children in our sample.[11] For each academic year, we asked whether a child was enrolled and coded as 1 if they were enrolled and as 0 if they were not enrolled. For each academic year, we also asked how many months a child was enrolled. The response here ranges from 0 to 10 months.[12]

We use this information to compute whether or not a child was ever enrolled between 2008 and 2013 in preschools and primary schools and the cumulative months of enrollment in each between 2008 and 2013. Preschool is defined as enrollment in kindergarten and playgroups, which are the two most common types of center-based services for young children before primary school in Indonesia. The summary statistics of rate and duration of enrollment are shown in Table 1. On average, 80 percent of girls report ever having attended preschools compared to 76 percent of boys. On average girls enroll for 14.3 months compared to 13.2 months for boys. Since all children were observed in primary school, there are no differences in enrollment rates though there is a small difference in duration of enrollment – on average girls have enrolled in primary schools for 0.5 months more than boys.

Quality of preschool services

The quality of preschool services in this paper is measured using the Early Childhood Environment Rating Scale (ECERS-R) (Harms, Clifford and Cryer, 2005). Two raters assessed

[9] The EDI domains have been shown to be correlated with children's readiness for school.
[10] EDI data were not collected for 9 year olds due to ceiling effects. There was very little variation at age 9 with almost all children scoring at the maximum end of the EDI scales, which range from 1 (low) to 10 (high).
[11] 2008 was chosen because it was the year before the Indonesia ECED project launched in 2009.
[12] The maximum number of months is 10 months since we follow the Indonesian academic calendar.

each center at the same time. Both raters were present in the room with the class they were observing for three hours and followed this group if they left the room for outdoor play. Raters did not interact with staff or students during their observation. The two raters scored each center on a seven-point Likert scale, which ranged from inadequate (1), minimal (3), good (5) to excellent (7). For each center, rater one and rater two's scores are averaged to construct a mean ECERS-R score.[13] We then computed village level averages of this ECERS-R score. In the analysis we divide the 310 villages in our sample in terciles – based on their average ECERS-R score. These are reported in Table 1 and show that girls and boys are equally likely to be found in villages with low, medium and high quality preschool centers.

Parenting practices

The primary caregivers of the children in our sample (usually mothers) were asked to answer a series of questions about their parenting practices. These practices were measured using 24 items describing parent-child relationships adapted from the Longitudinal Study of Australian Children (LSAC) (Zubrick et al. 2008). The questions covered a range of possible practices that reflect three domains: parental warmth, consistency, and hostility.

Caregivers were asked how often they used each of a number of different parenting practices. A total positive parenting score was given to each child's caregiver by adding together scores for each of the three parenting dimensions (with the negative items reversed). The higher the score, the more likely it is that parents have high levels of warmth and consistency, and low levels of hostility toward their children. These are reported in Table 1. On average, girls have a slightly higher total parenting practices score than boys but the magnitude of this difference is very small (0.91 points). In our analysis, we normalized this total parenting practices score to have a mean of 0 and a standard deviation of 1.

5. Gender gaps in early childhood development

We examine gender differences by comparing means between girls and boys by age group. Using ordinary least-squares (OLS), we regress each outcome of interest on gender. For

[13] These assessments focused on the seven subscales of the ECERS-R: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. All averages were done first by sub-scale and then overall to construct the center's ECERS-R score.

examining gender differences in preschool and primary school enrollment, we use a linear probability model.[14] Recall that our data are a cross-sectional sample of children between 6-9 years of age in 2013 and that we have their enrollment histories going back to 2008. Thus we when we report our estimates for enrollment, the age variable shows the age of the child at the time of enrollment. For instance the sample of 6-9 year olds in 2013 was 5-8 years old in 2012, 4-7 years old in 2011 and so on.

For ease of interpretation, we present the results of our estimations in a series of figures. Tables with point estimates are reported in the appendices.

5.1 Test scores

The results for overall test scores are presented in Figure 1 and there is a statistically significant gender gap favoring girls at all ages. At age 6, girls on average have a 0.245 SD higher score than boys. The female advantage over boys declines at age 7 to 0.125 SD, rises slightly to 0.144 SD at age 8, and then to 0.151 SD at age 9. Generally, the gender difference seems to be relatively consistent in the early years of primary school with girls having a statistically significant advantage over boys.

The results for both Bahasa Indonesia and mathematics components of the tests show similar results as the overall total score. Girls on average have a significantly higher score in Bahasa Indonesia than boys, ranging between 0.252 SD at age 6 to 0.160 SD at age 9. In mathematics, the female advantage is similar in magnitude to that found in the case of language, ranging from 0.243 SD at age 6 to 0.168 SD at age 9. In contrast, the gender gap in abstract reasoning is much smaller in magnitude compared to that observed for Bahasa Indonesia and mathematics. Girls have significantly higher scores than boys at age 6 (0.107 SD) but the gender gap disappears after that. This suggests that while girls may have an advantage over boys in Bahasa Indonesia and mathematics in the early years of primary school, they do not seem to have a similar advantage over boys in abstract reasoning.

[14] For enrollment in preschool, none were enrolled at age 9 (i.e., all were attending primary school). As a result, observations for children at age 9 (N=2,015) are omitted from the regressions on preschool enrollment. Similarly, for enrollment in primary school, all were enrolled at age 9 and none were enrolled at age 2 (N=1,428) and age 3 (N=4537) so these observations are omitted from the regressions on primary school enrollment.

5.2 Early Development Instrument (EDI)

Figure 2 presents gender gaps in each domain of the EDI for the sample of 6-8 year old children observed in 2013. Gender gaps in the EDI are statistically significant for the social competence, emotional maturity, and language and cognitive development domains at all ages. However, the magnitude of this difference varies across the domains. For social competence, the gender gap ranges from 0.166 SD at age 6 to 0.247 SD at age 8. Similarly for emotional maturity, the gender gap ranges between 0.279 SD at age 6 to and 0.296 SD at age 8. The gender gap in the language and cognitive domain is smaller: at age 6 it is 0.157 SD and shrinks down to 0.089 SD at age 8. For the physical health and well-being domain, there is no gender difference at age 6 and 7 but a slight statistically significant gender gap (0.068 SD) is apparent at age 8. Finally, for the communication skills and general knowledge domain, we also only observe a significant gender gap of 0.094 SD at age 8.

Thus, the results for test scores and EDI suggest the presence of gender gaps in both cognitive and non-cognitive skills in the first few years of primary school in rural Indonesia. At age 6, we already observe gender gaps in language and mathematics test scores as well as in children's social competence and emotional maturity. Given the existence of early gender gaps in rural Indonesia, we now examine whether there are gender differences in enrollment patterns and parenting practices to see if early schooling and family environment can be analyzed further as explanatory factors of the gender gap in test scores and EDI.

5.3 Education enrollment

For each child enrolled in primary school, we collected enrollment histories with regard to enrollment in education, including preschool education, starting from age two. These are used to construct data on enrollment by age. Two variables are constructed. The enrollment variable indicates whether or not the child was enrolled at any time, and the number of months of enrollment indicates the number of months the child was enrolled over a 12-month period.

Figure 3 reports the gender gaps and associated 95% confidence interval at each age for children in our sample both in terms of whether or not children are enrolled and in terms of duration of enrollment. At age 2, the earliest age for which we have enrollment histories, there is no observed gender gap. At age 3, girls are 4 percentage points more likely to be enrolled in preschool than boys, a point estimate that increases to 6.7 percentage points at age 4. At age 5 they are 5.7

percentage points more likely to enroll than boys. At age 6 there is no longer a gender gap in enrollment in early childhood education – a state that continues until age 8.

A gender gap is apparent in primary school enrollment as well. Girls are 5.4 percentage points more likely to be enrolled in primary school at age 6 than boys. At age 7, the gap is smaller – only 1.9 percentage points and there is no difference in enrollment rates when the children are age 8.

Months of enrollment differences between boys and girls follow a similar pattern to the rate of enrollment: girls are enrolled for more months in preschool on average at ages 3, 4, and 5 – though the point estimates range from 0.369 to 0.682 months. Likewise for enrollment in primary school – girls are enrolled for roughly 0.5 more months at age 6 but by age 7 there is virtually no difference in enrollment duration between boys and girls.

We assess whether these patterns vary at all by the average level of quality of early childhood education services. Figure 4 suggests that the gender gap in enrollment rates and duration is more pronounced in the villages that are in the highest tercile of quality. For instance, the gender gap at age 3 is 7.2 percentage points in the villages with the highest quality. In contrast, there is no enrollment gender gap in the villages with the lowest quality. For those children living in villages where quality is in the middle tercile, the gender gap is 3.7 percentage points – almost half of that observed in the highest quality tercile. A similar gradient along the quality dimension exists at age 4: 9.6 percentage points at the high end and 5 percentage points at the low end of the quality distribution. At age 5, the gradient of the gender gap with quality is less pronounced. Likewise Figure 4 suggests that in the case of months of enrollment there is a slightly bigger gender gap among children from villages with the highest quality and that this gender gap tends to diminish as we move from high to middle to low quality. However, the gap is not large – it tends to be one month or less at all ages.

5.4 Parenting score

The differences in parenting scores between girls and boys are presented in figure 5. As shown, parents of girls seem to exhibit more positive parenting behavior than parents of boys at ages 7 (0.114 SD), 8 (0.141 SD) and 9 (0.134 SD). The fairly consistent gender gaps in parenting score in the early years suggest the possibility of different parental expectations and behavior towards daughters and sons in Indonesia.

Thus, the data suggest there are statistically significant gender gaps in several early childhood outcomes in rural Indonesia. Specifically, girls perform better than boys on tests of Bahasa Indonesia and mathematics; and caregivers report higher levels of social competence, emotional maturity, and language and cognitive development for girls than for boys.

The data also reveal that girls are more likely to be enrolled in preschool at the appropriate age than boys and girls are cumulatively enrolled for more months of preschool than boys. Moreover, the gender gap in preschool enrollment is more pronounced in villages that (on average) have higher quality preschool services. In addition, there is evidence of gender differences in the family environment as parents of girls have higher parenting practices scores than parents of boys. Together, these patterns raise the question of whether the observed gender gaps in child development outcomes can be explained by gender differences in early schooling and family environments, as measured by gender differences in months of enrollment in preschool and primary school, the quality of preschool services and parenting score. In the next section, we explore this question further.

6. Decomposition of gender gaps by education enrollment and parenting score

6.1 Empirical model

Using an Oaxaca-Blinder decomposition, we investigate how much of the mean outcome difference between girls and boys is accounted for by gender differences in preschool and primary school enrollment and how much of it is accounted for by differences in parenting practices. The outcome variables we focus on are the standardized test scores and EDI in 2013. We distinguish between two types of predictors: decision variables and controls. Decision variables are factors influencing children's development that are decided by their parents and measure parental investments in the human capital of their children. In our analysis, the decision variables examined are (i) cumulative months in preschool between 2008 and 2013, (ii) cumulative months in primary school between 2008 and 2013 and (iii) total parenting score. In contrast, controls are factors influencing children's development that are not under direct control of their parents. Controls in our analysis are (i) education of mothers, (ii) household wealth and (iii) quality of preschools in the village.

The Oaxaca-Blinder decomposition is based on a linear regression model $Y = X'\beta + \varepsilon$ for girls and boys where Y is the outcome variable and X is a vector containing the predictor variables,

controls and a constant, β contains the slope and intercept parameters, and ε is the error term with $\mathrm{E}(\varepsilon) = 0$.

The mean outcome difference between girls and boys can be written as the difference in the linear prediction at the group-specific means of the predictors as follows:

$$\mathrm{E}(Y_{girls}) - \mathrm{E}(Y_{boys}) = E(X_{girls})'\beta_{girls} - E(X_{boys})'\beta_{boys}$$

because $\mathrm{E}(\varepsilon_{girls}) = 0$ and $\mathrm{E}(\varepsilon_{boys}) = 0$.

By rearranging this equation, we can identify the contribution of group differences in predictors to the overall outcome difference:

$$E(X_{girls})'\beta_{girls} - E(X_{boys})'\beta_{boys}$$
$$= \{E(X_{girls}) - E(X_{boys})\}'\beta_{all} + \{E(X_{girls})'(\beta_{girls} - \beta_{all}) + E(X_{boys})'(\beta_{all} - \beta_{boys})\}$$

where β_{all} is a vector of parameters from $Y_{all} = X_{all}'\beta_{all} + M\delta + \varepsilon$. M is an indicator variable equal to 1 for boys and 0 for girls.

Thus, the mean outcome difference between girls and boys has two components.[15] The first component $\{E(X_{girls}) - E(X_{boys})\}'\beta_{all}$ is the part of the outcome difference between girls and boys explained by group differences in the predictors. This first component is sometimes called the "endowment effect". The second component $\{E(X_{girls})'(\beta_{girls} - \beta_{all}) + E(X_{boys})'(\beta_{all} - \beta_{boys})\}$ is the "unexplained" part that captures all the potential effects of differences in other observed and unobserved characteristics between girls and boys. Our focus is on the endowment effect of the decision variables. This allows us to understand how much of the mean outcome difference is accounted for by group differences in months of enrollment in preschool and primary school as well as parenting practices.

The standard Oaxaca-Blinder decompositions, which we report in Tables 2, 4, 6 and 8, are based on the model described above, where decision variables and control variables are included separately in the model. One should however recognize that these decompositions provide an estimate of how much of the differences in development between boys and girls are associated with differences in human capital investments, rather than caused by differences in human capital investment. This is because our model is not a causal model. The investment in human capital of

[15] This is a modification of the original Oaxaca-Blinder decomposition, which has three components. See Jann 2008 for details.

children is determined by the same observed and unobserved variables that also influence child development directly and therefore should be considered endogenous.

We therefore implement an extended version of the decomposition model above based on the total program effect (TPE) technique introduced by Elbers and Gunning (2014). The TPE is a regression method in the program evaluation literature that deals with treatment heterogeneity. Instead of estimating average impact coefficients for each intervention of a program that consists of multiple interventions, the TPE method estimates the expected value of the total impact of the combined interventions for different subgroups of the population. The technique is important when there is considerable variation in the extent of exposure to each of the interventions. In practice, the TPE technique involves regressing the outcome of interest on a vector of the intervention variables, a vector of observed controls, and interactions of each intervention variable with each control variable.[16] Then, all terms involving the intervention variables are summed to calculate the total program effect. The TPE assumes that the way multiple interventions are assigned is in itself a characteristic of the program; thus, it measures the effect of the program inclusive of selectivity in the assignment of program interventions (see Elbers and Gunning 2014 for details and proofs).

The results of this extension of the standard Oaxaca-Blinder decompositions are in Tables 3, 5, 7, and 9.

$$Y_{all} = I_{all}\beta_1 + I_{all}\otimes Z_{all}\beta_2 + Z_{all}\beta_3 + M\delta + \varepsilon$$

where *I* now includes the decision variables and *Z* the controls. The additive specification and the linear effect that is assumed for the intervention reduces the computational complexity compared with propensity score matching methods. The endowment effect for the entire population is the weighted sum of the endowment effects of each of the subgroups defined by *Z*. This can be calculated as:

$$\{E(I_{girls}) - E(I_{boys})\}'\beta_1 + \sum_j \{E(IZ_{j,girls}) - E(IZ_{j,boys})\}'\beta_{j2}$$

where *j* identifies the control variable.

[16] Doing so allows us to address a key limitation of the standard decomposition which assumes that the relationship between the gender differences in the decision variables and the gender gap in cognitive and non-cognitive skills is the same for different levels of the controls. However, it is likely that control variables such as mother's education or household wealth moderate this relationship. For example, gender differences in months of preschool may have a small or large role to play in explaining the gender gaps in math test scores depending on mother's education level.

6.2 Results

The results of the decomposition analyses are presented in Tables 2-9. In these tables, the mean difference of standardized test scores or EDI scores between girls and boys is decomposed in differences resulting from gender differences in the cumulative months in preschool, the gender differences in the cumulative months in primary school, as well as the gender differences in parenting scores. The rest is unexplained and not reported.

For primary school test scores (shown in Tables 2 and 3), we analyze the scores on the language, mathematics and abstract reasoning by running separate regressions for each age. As noted earlier, the gender differences in test scores are concentrated in language and mathematics. For language, the gender difference is in the range of 0.26 to 0.15 SD and tends to reduce with age. For mathematics, the decline in test score with age is less pronounced, with girls outperforming boys by 0.23 SD at age 6 and 0.18 SD at age 9.

The results are very similar across the two methods. We find that for children age 6 and 7, preschool enrollment is the most important decision variable and explains around 0.035 to 0.029 SD, which is around 14 percent of the gender gap. For children age 8 and 9, gender differences in primary school enrollment become more important. In language, gender differences in primary school enrollment explain 18 and 29 percent of the gender gap for children at ages 8 and 9 respectively. In math, the proportion of the gender gap explained by duration of primary school enrollment is 12 to 20 percent.

However, we find slightly different results for parenting between the standard Oaxaca-Blinder decomposition and the extended model. The standard decomposition in Table 2 shows a few statistically significant coefficients for parenting score on gender gaps in language and math scores. It is worth noting, however, that the magnitude of the association is quite small, ranging from 0.004 to 0.006 SD, or up to approximately 4 percent of the gender gaps. In contrast, the extended decomposition in Table 3 does not yield any statistically significant effects for parenting score. This suggests that the effect of parenting score varies by the level of control variables (mother's education, household wealth, and preschool quality). Given that the improving preschool quality is particularly amenable to policy intervention, we now turn to examining how preschool quality moderates the relationship between decision variables and test scores in explaining the gender gap.

Tables 4 and 5 present the decomposition results separately for children living in villages with high, medium, and low quality preschool services. Overall, the results are similar for the two methods. The fraction of the gender gap in language and mathematics that is explained by gender differences in preschool enrollment increases with the quality of services. In areas with low quality, gender differences in enrollment in early childhood education have no statistically significant contribution to the gender gap, even though for most ages, the gender gap is higher than in areas with higher preschool quality. The contribution of primary school enrollment on gender gaps in language and math scores is similar in magnitude as the previous results in Tables 2 and 3, with the contribution concentrated at age 8 and 9. However, the standard errors in Tables 4 and 5 are larger than those of Tables 2 and 3, resulting in less significant effects. The effects of primary school are found across villages with high, medium and low quality preschool services. As before, we find small, significant coefficients of gender differences in parenting score in the standard decomposition but no significant contribution in the extended decomposition model.

For EDI scores, we also analyze each of the five domains separately by age. We focus our analysis on the social competence, emotional maturity, and language and cognitive development domains given that these are the three areas in which our previous descriptive analyses showed consistently significant gender differences. The results of the decompositions are quite different depending on which method is used. A general pattern we find is that gender differences in parenting explain a significant proportion of the gaps for all domains of the EDI if the standard decomposition is used, whereas no contribution is found if the extended version is used. This suggests that the gender based differences in parenting practices are correlated with the control variables. Conditioning on the control variables removes the effect of gender differences in parenting. Using the standard decomposition method (Table 6), we find significant contributions of gender differences in preschool enrollment for all domains for children age 7 and 8. For language and cognitive development, the contribution of preschool enrollment is quite small at about 4 percent, but for the other domains it ranges in the order of 15 to 29 percent. The effects of gender differences in preschool enrollment are less pronounced if the extended method is used (see Table 7). The effects particularly weaken for the 6 year old children in social competence. For the 7 year olds, gender differences in preschool enrollment contribute 21and 20 percent to the gender differences found in social competence and language and cognitive development respectively, regardless of the method used. For primary school enrollment we find significant

effects across all domains for children at age 8; around 21 percent for the language and cognitive domain and 3 to 6 percent for the social competence and emotional maturity domains. The results for primary school enrollment are rather similar across methods.

As before, we also unpack our EDI analysis into children from villages of high, medium, and low quality preschools. These results are presented in Tables 8 and 9. Once again, we focus on the EDI domains where we observe significant gender gaps, which are social competence, emotional maturity, and language and cognitive development. We find that gender differences in preschool enrollment explain gender differences in social competence and language and cognitive development in areas with high quality preschool if the standard decomposition method is used, but the effects become mostly insignificant if the extended decomposition method is used. This suggests that children's background characteristics (control variables) correlate with the quality of preschool services. In the extended decomposition model, the effects of gender differences in primary school enrollment are observed at age 8, and are mostly visible for the social competence and language and cognitive domains. An interesting pattern occurs for the parenting variables. Using the standard method, gender differences in parenting are important for the gender differences in the social competence and emotional maturity domains of the EDI for children living in villages with high quality preschools, and for children living in villages with low quality preschools. The significant coefficient on parenting is also observed in the extended model (Table 9), although the effect is weaker. It is not clear what drives this pattern.

7. Discussion and Conclusion

We find consistent gender gaps in total test scores at all ages, with substantial female advantage. The gender gaps in total test scores are driven by differences in children's performance in the language and mathematics portion of the test, but not by differences in performance in the abstract reasoning section of the assessment. Our findings here are consistent with previous studies. Girls outperform boys in language as early as age 6 (by 0.252 SD), which is similar to the 0.162 SD female advantage observed in the U.S. among kindergarteners (Cornwell et al. 2012, DiPrete and Jennings 2012). However, our results diverge from earlier studies that have observed girls losing ground to boys in mathematics during primary school, both in high-income countries (Fryer and Levitt 2010) and in lower- and middle-income countries (Bhardwaj et al. 2015, Dickerson et al. 2013). Instead, we find a female advantage in mathematics during the first few years of

schooling, which is consistent with an earlier study from Indonesia showing female advantage in mathematics at age 11 (Suryadarma 2015).

In addition, we find gender gaps in three out of five EDI domains: social competence, emotional maturity, and language and cognitive development. In all cases, we find evidence of female advantage. In the physical health and well-being and communication skills and general knowledge domains, gender differences are not consistently observed across ages 6 to 8. This is consistent with other studies using the EDI, which have found that girls score higher on the EDI in at least two of these domains in kindergarten (Australian Government 2013, Janus and Duku 2007).

The magnitude of the gender gap also ranges between the EDI domains, with the largest gender gap found in the emotional maturity domain followed by the social competence domain. This too is consistent with previous studies of the EDI in higher-income contexts, which have found that girls tend to develop faster than boys in these two domains (Australian Government, 2013, Janus and Duku 2007). The gender gaps in EDI observed in our study are somewhat large compared to results from other gender gap studies that use the EDI. Studies from Australia and Canada using the EDI show that boys have about a 0.09 standard deviation lower score than girls (Australian Government, 2013, Janus and Duku 2007). In contrast, we find gender gaps as large as 0.247 SD in the social competence domain and as large as 0.296 SD in the emotional maturity domain.

Our findings also show that there are significant gender gaps in enrollment both in terms of rate and duration in early childhood education and primary school. The magnitude of this gap varies by the age of the children. For preschool, both the rate and duration of enrollment for girls is higher than for boys in the prime preschool-going ages of 3 to 5. In primary school, gender gaps in enrollment only exist at ages 6 and 7 and close subsequently. The magnitude of these gaps is larger when average quality of preschool services in the village is higher. Taken together, these results suggest that girls are more likely to be enrolled in early childhood education programs at the intended age. This suggests that girls are likely to benefit more from early childhood education programs because they are exposed to better quality, age appropriate developmental activities. Suryadarma (2015) presents evidence to support that the Indonesian school system favors girls to boys. In addition, gender gaps are observed in parenting practices for children age 7 to 9,

suggesting that there are also some gender differences in the parent-child interaction at play in rural Indonesia.

In our decomposition analyses, we explored the extent to which these observed gender gaps in test scores and EDI domains are in part explained by one's early schooling environment (gender differences in duration of preschool and primary school enrollment) and in part explained by one's family environment (gender differences in parenting practices). The standard Oaxaca-Blinder decomposition model revealed that preschool and primary school enrollment are more important for explaining the gender gap in test scores (measures of cognitive skills) than for explaining the gender gap in the social competence domain and emotional maturity domain of the EDI (measures of non-cognitive skills). The reverse is true for parenting scores, which are more important in explaining the gender gaps in non-cognitive skills than in explaining the gender gaps in cognitive outcomes.

However, the extended decomposition method – which takes into account how different levels of control variables may moderate the relationship between explanatory variables and outcome variables – suggests that on average, gender gaps in parenting scores are not significant predictors of either cognitive or non-cognitive skills. This means that parenting practices have no effect on child outcomes if the effect is conditioned on the mother's level of education, household wealth, and preschool quality.

Examining whether there is variation in the decomposition of gender gaps in cognitive and non-cognitive skills by the quality of preschool services available in the village reveals interesting contrasts. In both our standard and extended decomposition models, we find that gender differences in enrollment in preschool seem to matter most for test scores when the quality of available early childhood education services is high. This is not the case when quality is medium or low. Under the preferred decomposition method, gender differences in parenting practices do not explain gaps in test scores. However, parenting practices seem to matter for children's non-cognitive skills (social competence and emotional maturity) -- especially when the quality of available preschool services is high or low. This suggests that preschool quality and parenting simultaneously play important roles in children's early development and the emerging gender gap in cognitive and non-cognitive skills. We also find that duration of enrollment in preschool and primary school is not as important as parenting score in explaining the gender gap in non-cognitive skills.

This paper contributes to the gender gap literature by examining whether gender differences in cognitive and non-cognitive skills emerge in the first few years of primary school. Using data from rural Indonesia, we document the early emergence of gender gaps and show that a combination of early schooling and family environment explain the observed difference between girls and boys. From a policy standpoint, these early-emerging gender differences in rural Indonesia highlight the important role both schools and families play in equally supporting the needs of girls and boys starting from the early years. Differences in the magnitude of these roles at various levels of quality underscore the importance of ensuring the quality of education services – starting with preschools.

References

Anderson, M.L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484), 1481-1495.

Arteaga, I., Humpage, S., Reynolds, A. J., & Temple, J. A. (2014). One year of preschool or two: Is it important for adult outcomes? *Economics of Education Review*, *40*, 221-237.

Australian Government. (2013). *A snapshot of early childhood development in Australia 2012 - Aedi National Report.* Canberra: Australian Government.

Autor, D., Figlio, D., Karbownik, K., Roth, J. & Wasserman, M. (2016). Family disadvantage and the gender gap in behavioral and educational outcomes (No.22267). *National Bureau of Economic Research Working Paper.*

Baker, M., & Milligan, K. (2013). Boy-girl differences in parental time investments: Evidence from three countries (No. w18893). *National Bureau of Economic Research Working Paper*.

Bertrand, M., & Pan, J. (2011). The trouble with boys: Social influences and the gender gap in disruptive behavior (No. w17541). *National Bureau of Economic Research Working Paper*.

Bharadwaj, P., Dahl, G. B., & Sheth, K. (2014). Gender Discrimination in the Family. *The Economics of the Family: How the Household Affects Markets and Economic Growth*, *237*.

Bhardwaj, P., De Giorgi, G., Hansen, G. & Neilson, C. (2015). The gender gap in mathematics: Evidence from a middle-income country. *Federal Reserve Bank of New York Staff Report, No. 721*.

Brinkman, S., Kinnell, A., Maika, A., Hasan, A., Jung, H., & Pradhan, M. (2015). Validity and reliability of the Early Development Instrument in Indonesia. *Working Paper*.

Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Non-Cognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. Journal of Human Resources, 48(1), 236-264.

Cueto, S., Leon, J., Guerrero, G., & Muñoz, I. (2009). Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives. *Young Lives Technical Note*, *15*.

Cunha, F. & Heckman, J.J. (2008). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *The Journal of Human Resources*, 43(4), 738-782.

Cunha, F., Heckman, J.J. & Schennach, S.M. (2010). Estimating the technology of cognitive and non-cognitive skill formation. *Econometrica*, 78(3), 883-931.

Currie, J. and Thomas, D. (1999). Early test scores, socioeconomic status and future outcomes (No. w6943). *National Bureau of Economic Research Working Paper*.

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.

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.

Domitrovich, C. E., Morgan, N. R., Moore, J. E., Cooper, B. R., Shah, H. K., Jacobson, L., & Greenberg, M. T. (2013). One versus two years: Does length of exposure to an enhanced preschool program impact the academic functioning of disadvantaged children in kindergarten?. *Early Childhood Research Quarterly*, *28*(4), 704-713.

Engle, P. L., Fernald, L. C., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., ... & Iltus, S. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, *378*(9799), 1339-1353.

Forget-Dubois, N., Lemelin, J. P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education & Development*, 18(3), 405-426.

Fryer, R.G. & Levitt, S.D. (2010). Empirical Analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics 2*, 210-240.

Goldin, C. (2006). *The quiet revolution that transformed women's employment, education, and family* (No. w11953). National Bureau of Economic Research.

Harms, T., Clifford, R. M., and Cryer, D. (2005). Early Childhood Environment Rating Scale - Revised Edition. New York: Teachers College Press.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453-479.

Janus, M. & E. Duku. (2007). The school entry gap: Socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education & Development*, 18(3), 375-403.

Janus, M. & D. Offord**.** (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science*, 39, 1-22.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, *26*(1), 52-66.

Mullis, I.V.S, Martin, M.O., Foy,P., & Drucker, K.T**.** (2012). *PIRLS 2011 International Results in Reading*. Chestnut Hill, MA: International Association for the Evaluation of Academic Achievement (IEA).

Nores, M., & Barnett, W. S. (2010). Benefits of early childhood interventions across the world:(Under) Investing in the very young. *Economics of education review*, *29*(2), 271-282.

OECD**.** (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence.* PISA, OECD Publishing.

Owens, J. (2013). Habits that make, habits that break: Early childhood behavior problems and the gender gap in education in the United States. *Paper presented at Population Association of America 2014*.

Paxson, C. H. & Schady, N. R. (2005). *Cognitive development among young children in Ecuador: The roles of wealth, health and parenting* (Vol. 3605). World Bank Publications.

Pradhan, M., Brinkman, S. A., Beatty, A., Maika, A., Satriawan, E., de Ree, J., & Hasan, A. (2013). Evaluating a community-based early childhood education and development program in Indonesia: study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group. *Trials*, 14(1), 259.

Suryadarma, D. (2015). Gender differences in numeracy in Indonesia: evidence from a longitudinal dataset. *Education Economics*, 23(2), 180-198.

World Bank. (2011). *World Development Report 2012: Gender equality and development.* Washington D.C.: World Bank.

World Bank (2015). *World Development Indicators.*Washington D.C.: World Bank.

Zubrick, S., G. J. Smith, J. M. Nicholson, A. V. Sanson, and T. A. Jackiewicz. 2008. Parenting and Families in Australia . Canberra: FaHCSIA (Department of Families, Housing, Community Services and Indigenous Affairs).

Tables and Figures

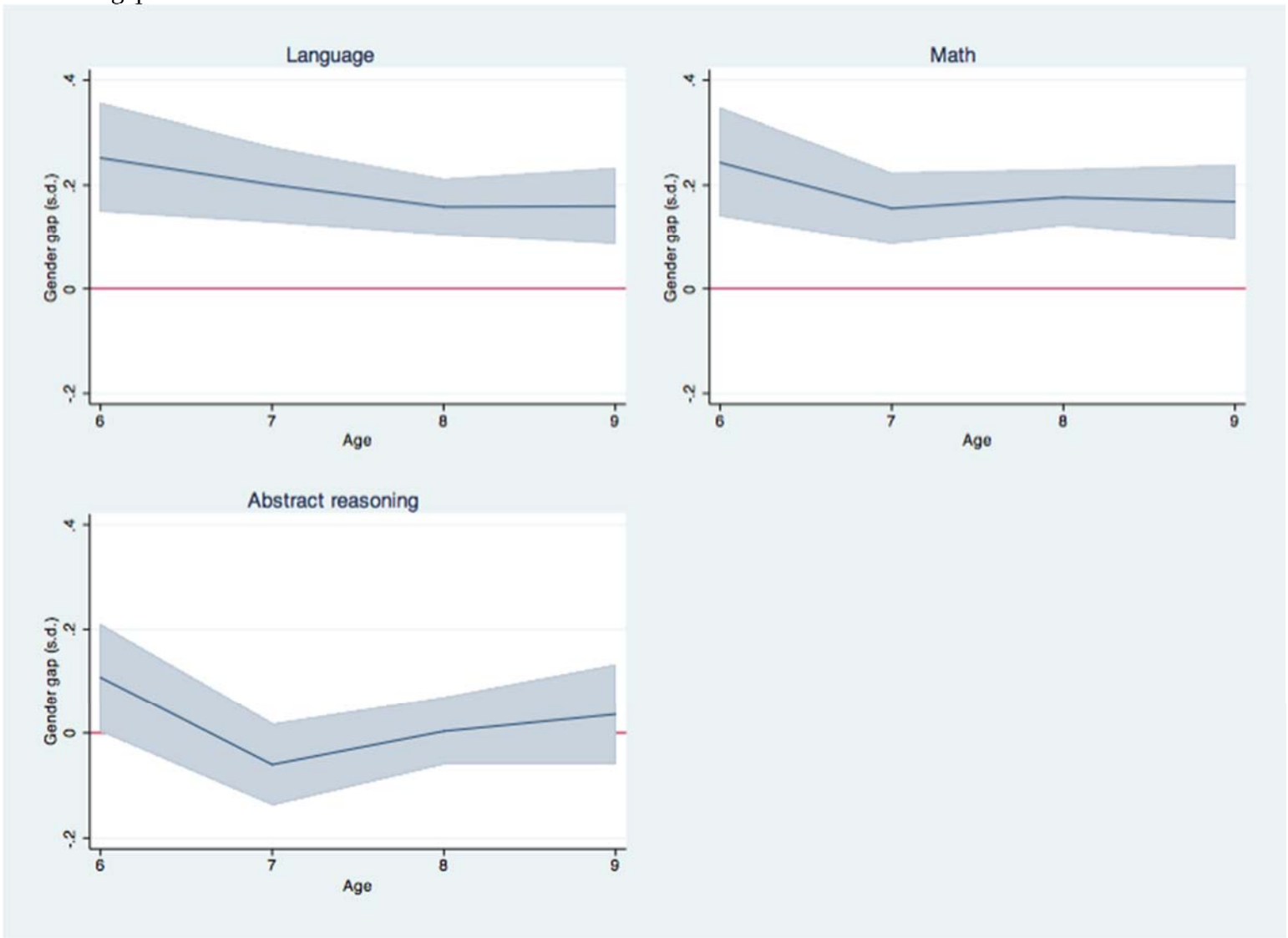
Table 1. Summary statistics of key variables

	Girls (N=5380)				Boys (N=5478)				Gender difference (Girls - Boys)	
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max	Est.	S.E.
Test score (raw scores)										
Language	6.491	2.865	0	10	5.975	2.949	0	10	0.52***	(0.06)
Math	10.196	4.374	0	15	9.400	4.502	0	15	0.80***	(0.09)
Abstract reasoning	5.947	3.274	0	14	5.945	3.442	0	14	0.00	(0.06)
Early Development Instrument (only for age 8 and younger)										
Physical health & well-being	9.120	1.096	2.5	10	9.063	1.152	3.333	10	0.06**	(0.02)
Social competence	7.692	1.378	3.75	10	7.449	1.389	0	10	0.24***	(0.03)
Emotional maturity	7.065	1.347	0.833	10	6.684	1.407	2.083	10	0.38***	(0.03)
Language & cognitive development	9.387	1.421	0	10	9.094	1.677	0	10	0.29***	(0.03)
Communication skills & general knowledge	7.680	2.180	0	10	7.526	2.210	0	10	0.15***	(0.05)
Ever enrolled in preschool (Yes =1)	0.802	0.399	0	1	0.760	0.427	0	1	0.04***	(0.01)
Ever enrolled in primary (Yes =1)	0.998	0.049	0	1	0.997	0.052	0	1	0.00	(0.00)
Cumulative months in preschool (between ‘08-‘13)	14.317	9.855	0	40	13.202	9.857	0	40	1.11***	(0.19)
Cumulative months in primary (between ‘08-‘13)	19.097	9.433	0	48	18.571	9.187	0	48	0.53***	(0.18)
Parenting score	81.003	7.279	56	109	80.090	7.409	45	103	0.91***	(0.14)
Low preschool quality (Yes=1)	0.325	0.468	0	1	0.309	0.462	0	1	0.02*	(0.01)
Medium preschool quality (Yes=1)	0.327	0.469	0	1	0.339	0.473	0	1	-0.01	(0.01)
High preschool quality (Yes=1)	0.348	0.476	0	1	0.352	0.478	0	1	-0.00	(0.01)
Age	7.625	0.937	6	9	7.647	0.924	6	9	-0.02	(0.02)
Mother's education (years)	7.305	3.691	0	15	7.349	3.636	0	15	-0.04	(0.07)
Household wealth (z-score)	0.079	0.936	-3.573	2.216	0.090	0.940	-3.531	2.248	-0.01	(0.02)

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

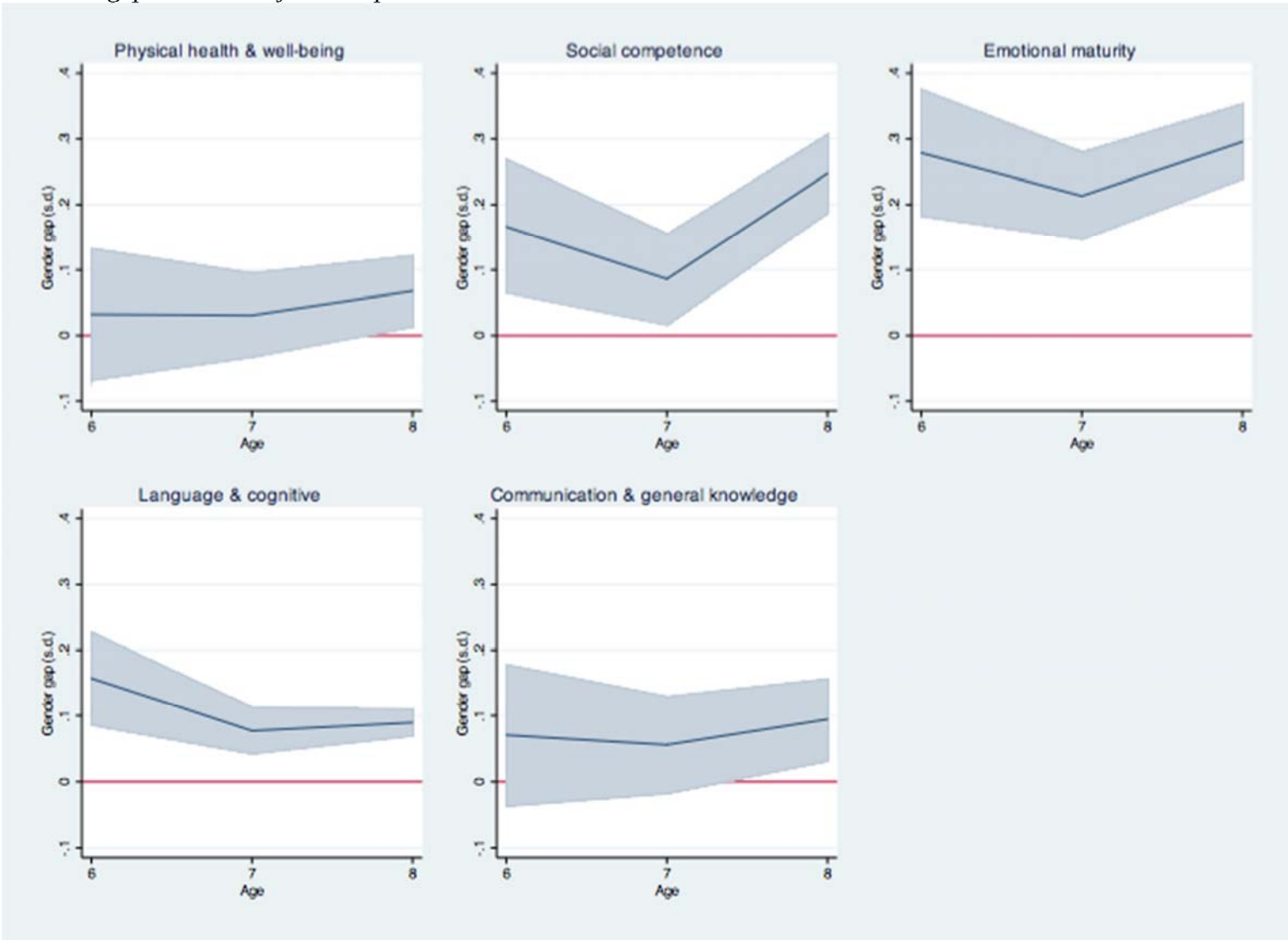
Note: All variables measured in 2013. For EDI, the sample size is reduced to 8653 children (4309 girls and 4344 boys) since 9 year-olds are not included.

Figure 1. Gender gaps in test scores



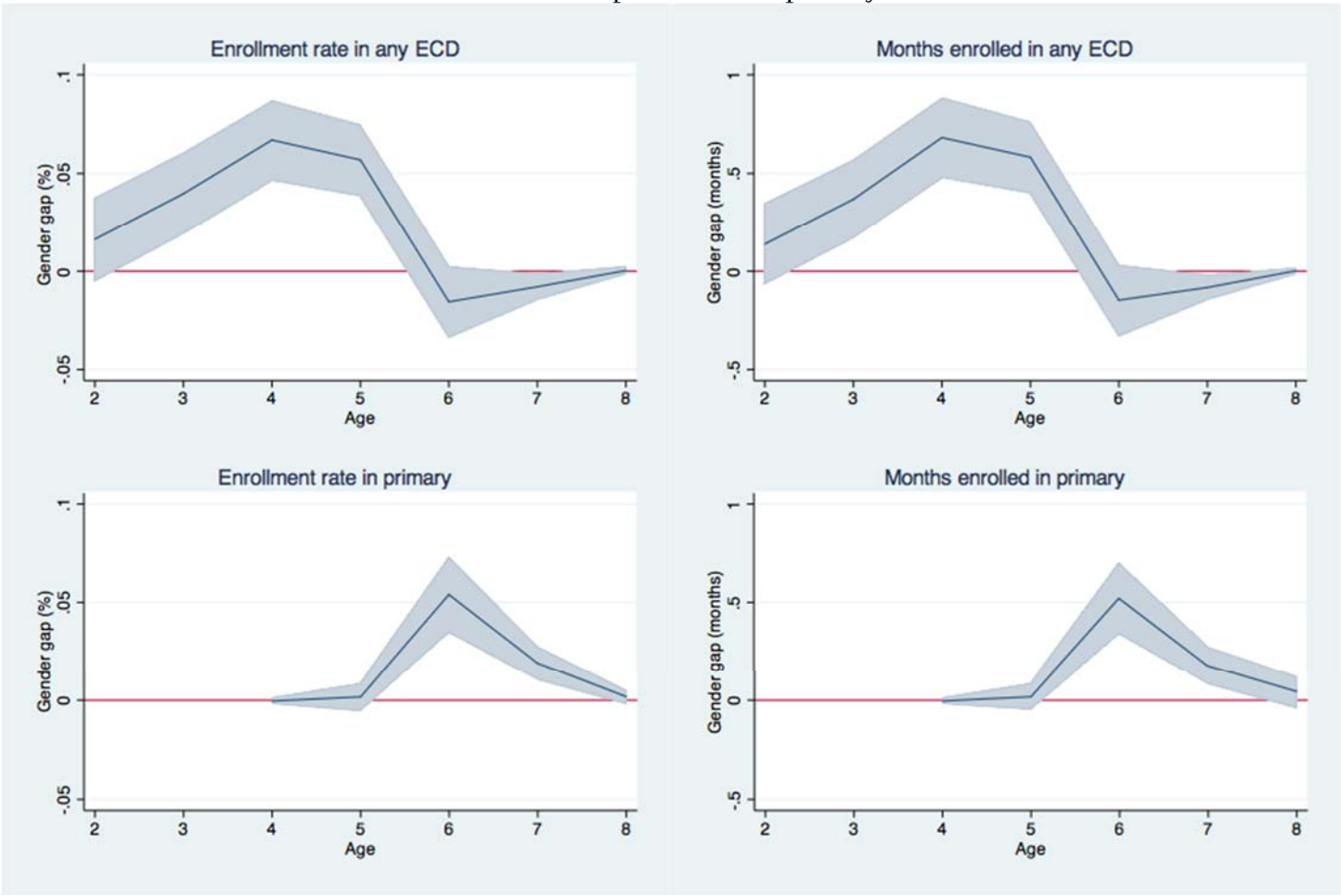
Note: Positive gender gap means that the predicted outcomes for girls are higher than that of boys. Point estimates and 95% confidence interval shown. See Appendix A1 for the regression estimates from which these figures are drawn.

Figure 2. Gender gaps in the Early Development Instrument



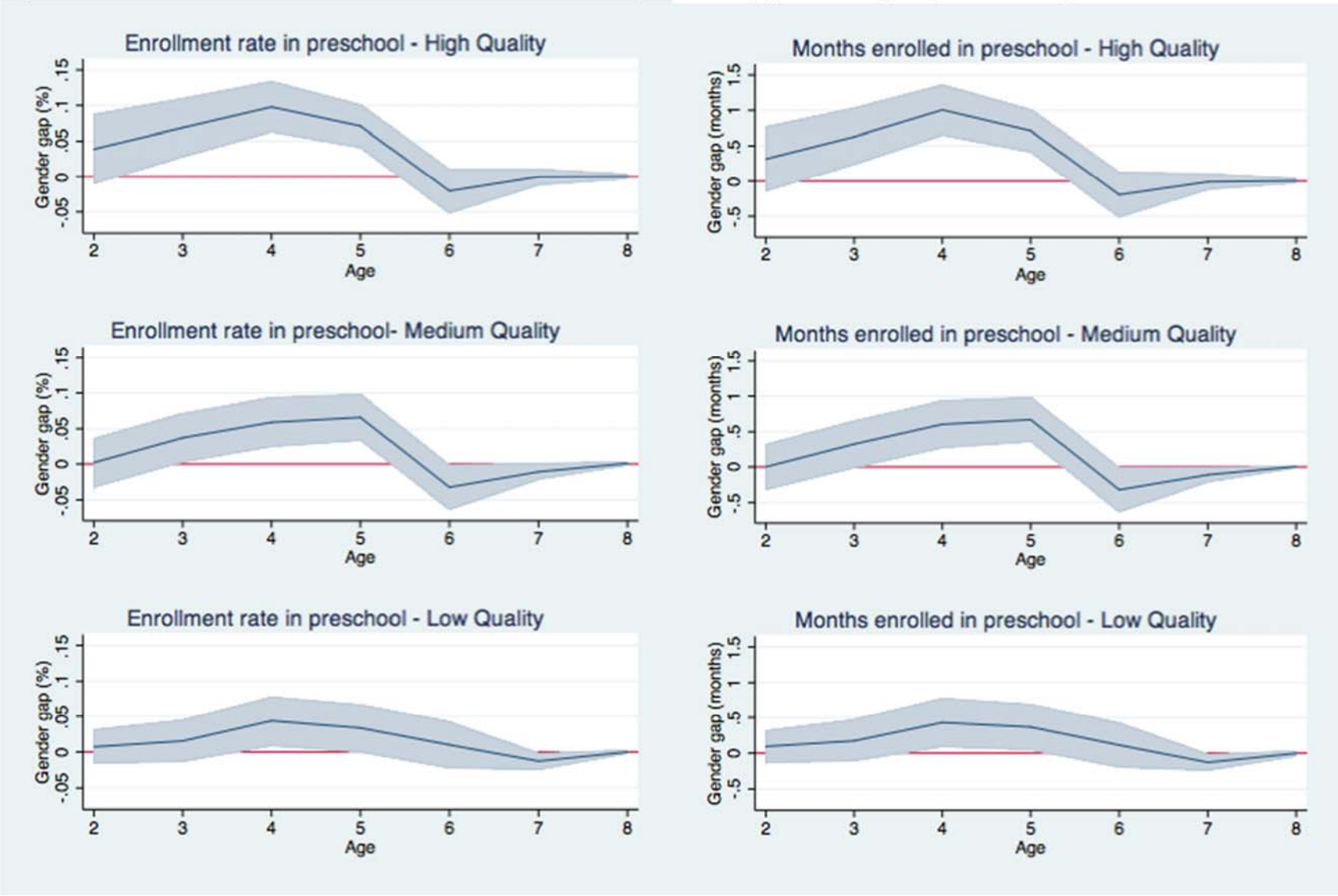
Note: Positive gender gap means that the predicted outcomes for girls are higher than that of boys. Point estimates and 95% confidence interval shown. See Appendix A2 for the regression estimates from which these figures are drawn.

Figure 3: Gender differences in rate and enrollment months in preschool and primary school



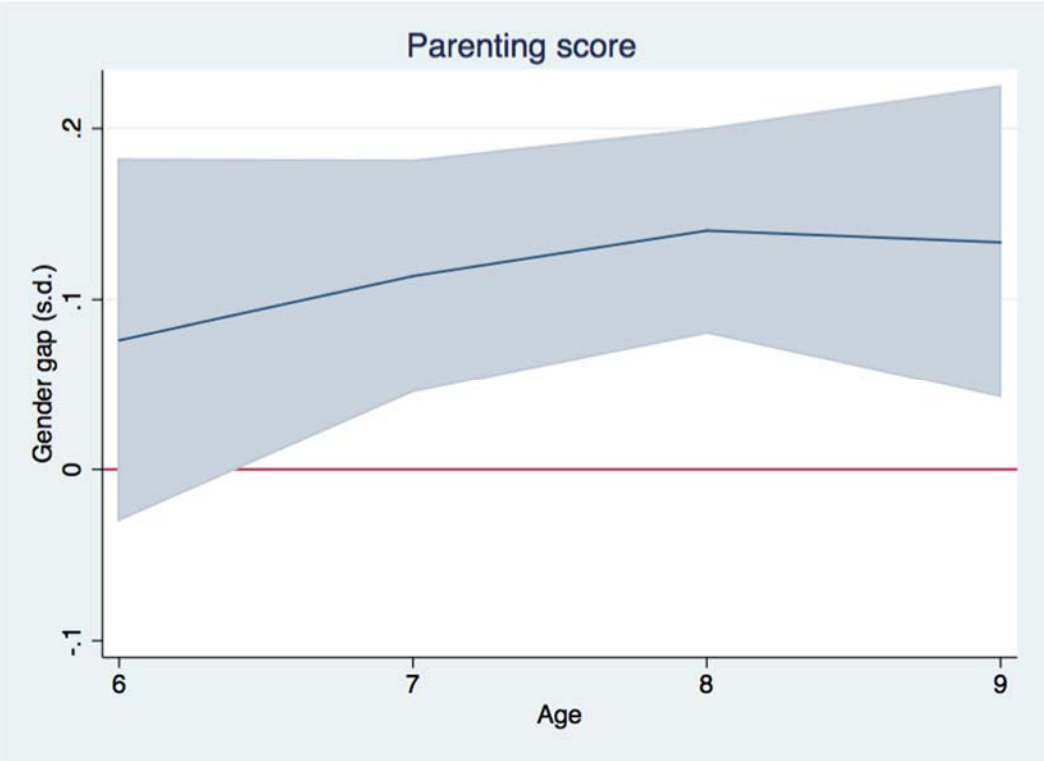
Note: Point estimates and 95% confidence interval shown. Positive gender gap means that the predicted outcomes for girls are higher than that of boys. Graphs are constructed using enrollment histories collected in 2013 from a sample of 6 to 9 year-old children. Histories go back to the 2008-2009 academic year when 6 year olds in 2012-2013 would have been 2 years old in 2008-2009, and 9 year-olds in 2012-2013 would have been 5 years old in 2008-2009. In the graph, age 9 is dropped because it perfectly predicts outcomes: all children are enrolled in primary school and none are enrolled in preschool. Age 2 and 3 are dropped for enrollment in primary school because they perfectly predict outcomes: no child that young is enrolled in primary school. See Appendix A3 for the regression estimates from which these figures are drawn.

Figure 4. Gender differences in rate and enrollment months in preschool by preschool quality at the village-level



Note: Positive gender gap means that the predicted outcomes for girls are higher than that of boys. Point estimates and 95% confidence interval shown. See Appendix A4 for the regression estimates from which these figures are drawn..

Figure 5. Gender gaps in parenting score



Note: Positive gender gap means that the predicted outcomes for girls are higher than that of boys. Point estimates and 95% confidence interval shown. See Appendix A5 for the regression estimates from which these figures are drawn.

Table 2. Gender gaps in test scores explained by differences in education enrollment and parenting (standard decomposition)

	Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score		Sample Size
		Est	SE	Est	SE	Est	SE	Est	SE	
Language	6	0.262***	(0.054)	0.035***	(0.013)	-0.000	(0.007)	0.007	(0.005)	1,323
	7	0.200***	(0.038)	0.027***	(0.006)	0.013	(0.009)	0.005**	(0.003)	2,869
	8	0.161***	(0.028)	0.008**	(0.004)	0.029***	(0.007)	0.004*	(0.002)	3,936
	9	0.157***	(0.038)	-0.001	(0.005)	0.046***	(0.011)	0.006*	(0.003)	1,838
Math	6	0.234***	(0.054)	0.034***	(0.012)	-0.000	(0.004)	0.004	(0.004)	1,323
	7	0.157***	(0.036)	0.022***	(0.006)	0.010	(0.007)	0.005**	(0.003)	2,869
	8	0.178***	(0.028)	0.008**	(0.004)	0.022***	(0.005)	0.004*	(0.002)	3,936
	9	0.179***	(0.038)	-0.001	(0.004)	0.035***	(0.009)	0.005	(0.003)	1,838
Abstract reasoning	6	0.097*	(0.055)	0.018**	(0.008)	-0.000	(0.004)	0.000	(0.002)	1,323
	7	-0.060	(0.041)	0.014***	(0.005)	0.009	(0.006)	-0.000	(0.002)	2,869
	8	0.006	(0.034)	0.004*	(0.002)	0.021***	(0.005)	0.005**	(0.002)	3,936
	9	0.047	(0.051)	-0.000	(0.002)	0.031***	(0.009)	0.009**	(0.005)	1,838

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

Note: Each regression also includes controls. Control variables are preschool quality, mother's education in years, and household wealth (*z*-score)

Table 3. Gender gaps in test scores explained by differences in education enrollment and parenting (extended decomposition)

	Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score		Sample Size
		Est	SE	Est	SE	Est	SE	Est	SE	
Language	6	0.262***	(0.054)	0.032**	(0.014)	0.000	(0.014)	-0.011	(0.051)	1,323
	7	0.200***	(0.038)	0.028***	(0.010)	0.012	(0.009)	0.012	(0.011)	2,869
	8	0.161***	(0.028)	0.009*	(0.006)	0.029***	(0.007)	0.000	(0.008)	3,936
	9	0.157***	(0.038)	-0.002	(0.006)	0.043***	(0.013)	0.008	(0.025)	1,838
Math	6	0.234***	(0.054)	0.034**	(0.015)	-0.002	(0.008)	-0.016	(0.032)	1,323
	7	0.157***	(0.036)	0.022***	(0.008)	0.009	(0.008)	-0.004	(0.012)	2,869
	8	0.178***	(0.028)	0.010**	(0.005)	0.020***	(0.007)	0.002	(0.008)	3,936
	9	0.179***	(0.038)	-0.003	(0.006)	0.030**	(0.013)	-0.012	(0.023)	1,838
Abstract reasoning	6	0.097*	(0.055)	0.023**	(0.010)	-0.011	(0.011)	-0.030	(0.032)	1,323
	7	-0.060	(0.041)	0.013	(0.009)	0.014	(0.009)	0.004	(0.022)	2,869
	8	0.006	(0.034)	0.003	(0.004)	0.022***	(0.007)	0.002	(0.011)	3,936
	9	0.047	(0.051)	-0.003	(0.004)	0.032***	(0.012)	0.004	(0.017)	1,838

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

Note: Each regression also includes controls. Control variables are center quality, mother's education in years, and household wealth (*z*-score). The extended decomposition model also interacts each control variable with each decision variable (cumulative months of enrollment in preschool, cumulative months in primary, and total parenting score).

Table 4. Gender gaps in test scores explained by difference in education enrollment and parenting by preschool quality (standard decomposition)

		Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score	
			Est	SE	Est	SE	Est	SE	Est	SE
High quality preschool	Language	6	0.134	(0.099)	0.093***	(0.035)	-0.005	(0.008)	-0.006	(0.016)
		7	0.285***	(0.068)	0.065***	(0.018)	0.015	(0.017)	0.004	(0.005)
		8	0.104**	(0.046)	0.017*	(0.009)	0.026**	(0.011)	0.004	(0.004)
		9	0.068	(0.065)	-0.009	(0.011)	0.043**	(0.022)	0.013*	(0.008)
	Math	6	0.167*	(0.098)	0.082***	(0.032)	-0.002	(0.005)	-0.003	(0.007)
		7	0.247***	(0.064)	0.047***	(0.014)	0.011	(0.013)	0.011*	(0.007)
		8	0.113**	(0.048)	0.017*	(0.010)	0.022**	(0.010)	0.004	(0.004)
		9	0.084	(0.064)	-0.009	(0.011)	0.034*	(0.018)	0.016**	(0.008)
	Abstract reasoning	6	0.055	(0.098)	0.023	(0.016)	-0.005	(0.007)	-0.000	(0.002)
		7	0.029	(0.071)	0.036***	(0.013)	0.008	(0.009)	-0.003	(0.005)
		8	0.057	(0.057)	0.009	(0.006)	0.015**	(0.007)	0.001	(0.005)
		9	0.069	(0.084)	-0.003	(0.004)	0.021*	(0.012)	0.020*	(0.011)
Medium quality preschool	Language	6	0.243***	(0.093)	0.004	(0.015)	0.008	(0.016)	0.002	(0.004)
		7	0.113*	(0.065)	0.039***	(0.014)	0.030*	(0.018)	0.001	(0.004)
		8	0.193***	(0.048)	0.010	(0.006)	0.033***	(0.012)	0.002	(0.002)
		9	0.141**	(0.064)	0.002	(0.006)	0.040**	(0.016)	-0.001	(0.003)
	Math	6	0.208**	(0.092)	0.005	(0.019)	0.004	(0.008)	0.002	(0.004)
		7	0.103*	(0.061)	0.037***	(0.013)	0.023*	(0.014)	0.001	(0.003)
		8	0.237***	(0.048)	0.008	(0.005)	0.025***	(0.009)	0.004	(0.003)
		9	0.225***	(0.064)	0.001	(0.004)	0.028**	(0.013)	-0.001	(0.003)
	Abstract reasoning	6	-0.026	(0.089)	0.004	(0.015)	0.005	(0.010)	-0.001	(0.003)
		7	-0.136*	(0.071)	0.029**	(0.012)	0.022	(0.014)	0.000	(0.000)
		8	-0.040	(0.058)	0.005	(0.004)	0.030***	(0.011)	0.005	(0.004)
		9	-0.009	(0.093)	0.001	(0.002)	0.034**	(0.016)	-0.002	(0.004)
Low quality preschool	Language	6	0.412***	(0.089)	0.025	(0.016)	0.009	(0.015)	0.007	(0.010)
		7	0.207***	(0.063)	0.001	(0.003)	-0.001	(0.014)	0.009	(0.006)
		8	0.208***	(0.051)	0.001	(0.002)	0.027**	(0.013)	0.004	(0.004)
		9	0.293***	(0.069)	0.002	(0.005)	0.046**	(0.020)	0.008	(0.007)
	Math	6	0.331***	(0.093)	0.023	(0.015)	0.006	(0.011)	0.010	(0.011)
		7	0.131**	(0.062)	0.000	(0.003)	-0.001	(0.012)	0.003	(0.005)
		8	0.199***	(0.050)	0.003	(0.003)	0.019**	(0.010)	0.002	(0.003)
		9	0.254***	(0.068)	0.002	(0.005)	0.034**	(0.015)	0.000	(0.007)
	Abstract reasoning	6	0.266***	(0.095)	0.017	(0.013)	0.003	(0.005)	0.005	(0.012)
		7	-0.066	(0.069)	-0.005	(0.005)	-0.001	(0.010)	0.001	(0.005)
		8	0.010	(0.059)	0.001	(0.002)	0.019*	(0.010)	0.005	(0.004)
		9	0.090	(0.089)	0.002	(0.007)	0.038**	(0.018)	0.013	(0.010)

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

Note: Each regression also includes controls. Control variables include mother's education and household wealth (*z*-score).

Table 5. Gender gaps in test scores explained by difference in education enrollment and parenting by preschool quality (extended decomposition)

		Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score	
			Est	SE	Est	SE	Est	SE	Est	SE
High quality preschool	Language	6	0.134	(0.099)	0.095**	(0.042)	-0.022	(0.033)	-0.044	(0.089)
		7	0.285***	(0.068)	0.068***	(0.019)	0.024	(0.019)	-0.000	(0.045)
		8	0.104**	(0.046)	0.015*	(0.009)	0.025**	(0.011)	0.003	(0.025)
		9	0.068	(0.065)	-0.012	(0.013)	0.054*	(0.030)	-0.005	(0.039)
	Math	6	0.167*	(0.098)	0.101***	(0.039)	-0.003	(0.013)	-0.035	(0.085)
		7	0.247***	(0.064)	0.039***	(0.015)	0.012	(0.017)	0.022	(0.047)
		8	0.113**	(0.048)	0.016	(0.009)	0.019*	(0.011)	0.005	(0.016)
		9	0.084	(0.064)	-0.004	(0.013)	0.033	(0.021)	0.003	(0.033)
	Abstract reasoning	6	0.055	(0.098)	0.076**	(0.036)	-0.005	(0.014)	-0.102	(0.106)
		7	0.029	(0.071)	0.059***	(0.020)	0.026	(0.017)	0.009	(0.055)
		8	0.057	(0.057)	0.010	(0.007)	0.015	(0.010)	0.000	(0.022)
		9	0.069	(0.084)	0.003	(0.011)	0.002	(0.039)	0.018	(0.045)
Medium quality preschool	Language	6	0.243***	(0.093)	-0.001	(0.021)	-0.018	(0.030)	0.052	(0.090)
		7	0.113*	(0.065)	0.029	(0.018)	0.033*	(0.019)	0.042	(0.049)
		8	0.193***	(0.048)	0.009	(0.007)	0.034***	(0.013)	-0.004	(0.016)
		9	0.141**	(0.064)	-0.003	(0.013)	0.040*	(0.023)	0.054	(0.057)
	Math	6	0.208**	(0.092)	0.013	(0.023)	-0.007	(0.026)	-0.030	(0.087)
		7	0.103*	(0.061)	0.038**	(0.017)	0.025	(0.015)	0.059	(0.057)
		8	0.237***	(0.048)	0.009	(0.006)	0.024*	(0.013)	-0.005	(0.023)
		9	0.225***	(0.064)	-0.009	(0.014)	0.029	(0.018)	0.040	(0.051)
	Abstract reasoning	6	-0.026	(0.089)	-0.014	(0.026)	-0.041	(0.035)	-0.037	(0.090)
		7	-0.136*	(0.071)	0.027*	(0.015)	0.023	(0.015)	0.108	(0.079)
		8	-0.040	(0.058)	0.004	(0.006)	0.027*	(0.014)	0.004	(0.032)
		9	-0.009	(0.093)	-0.029	(0.019)	0.038	(0.024)	-0.061	(0.074)
Low quality preschool	Language	6	0.412***	(0.089)	0.010	(0.019)	0.005	(0.035)	-0.010	(0.094)
		7	0.207***	(0.063)	-0.001	(0.007)	-0.007	(0.016)	-0.004	(0.031)
		8	0.208***	(0.051)	-0.000	(0.004)	0.027**	(0.014)	0.003	(0.009)
		9	0.293***	(0.069)	0.005	(0.012)	0.058**	(0.029)	-0.035	(0.059)
	Math	6	0.331***	(0.093)	0.014	(0.017)	0.008	(0.023)	-0.018	(0.084)
		7	0.131**	(0.062)	-0.002	(0.005)	-0.003	(0.014)	0.018	(0.028)
		8	0.199***	(0.050)	0.004	(0.004)	0.019**	(0.010)	0.002	(0.014)
		9	0.254***	(0.068)	0.005	(0.010)	0.032	(0.033)	-0.037	(0.050)
	Abstract reasoning	6	0.266***	(0.095)	0.001	(0.021)	-0.005	(0.016)	-0.009	(0.054)
		7	-0.066	(0.069)	-0.003	(0.006)	0.003	(0.021)	0.027	(0.047)
		8	0.010	(0.059)	-0.002	(0.004)	0.019*	(0.011)	0.003	(0.018)
		9	0.090	(0.089)	0.005	(0.010)	0.024	(0.032)	-0.035	(0.067)

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

Note: Each regression also includes controls. Control variables include mother's education and household wealth (*z*-score). The extended decomposition model also interacts each control variable with each decision variable (cumulative months of enrollment in preschool, cumulative months in primary, and total parenting score).

Table 6. Gender gaps in EDI explained by difference in education enrollment and parenting (standard decomposition)

	Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score		Sample Size
		Est	SE	Est	SE	Est	SE	Est	SE	
Physical health & wellbeing	6	0.035	(0.054)	-0.004	(0.005)	0.001	(0.003)	0.011	(0.009)	1,321
	7	0.043	(0.034)	0.000	(0.003)	0.001	(0.001)	0.017***	(0.006)	2,859
	8	0.083***	(0.030)	0.002	(0.002)	0.006**	(0.003)	0.019***	(0.005)	3,801
Social competence	6	0.155***	(0.054)	0.016**	(0.007)	0.001	(0.002)	0.020	(0.015)	1,321
	7	0.090**	(0.037)	0.019***	(0.005)	0.007	(0.005)	0.026***	(0.009)	2,859
	8	0.260***	(0.032)	0.004*	(0.002)	0.014***	(0.004)	0.037***	(0.009)	3,801
Emotional maturity	6	0.272***	(0.052)	-0.004	(0.004)	0.001	(0.002)	0.022	(0.016)	1,321
	7	0.228***	(0.036)	-0.007*	(0.004)	0.001	(0.001)	0.031***	(0.011)	2,859
	8	0.306***	(0.031)	0.002	(0.001)	0.008***	(0.003)	0.041***	(0.010)	3,801
Language & cognitive skills	6	0.161***	(0.037)	0.021***	(0.008)	0.002	(0.004)	0.007	(0.006)	1,321
	7	0.082***	(0.019)	0.017***	(0.004)	0.009	(0.005)	0.003**	(0.001)	2,859
	8	0.089***	(0.011)	0.003*	(0.002)	0.015***	(0.003)	0.004***	(0.001)	3,801
Communication skills & general knowledge	6	0.058	(0.056)	0.020**	(0.008)	0.000	(0.001)	0.007	(0.006)	1,321
	7	0.045	(0.039)	0.021***	(0.006)	0.007	(0.005)	0.008**	(0.004)	2,859
	8	0.111***	(0.033)	0.006*	(0.003)	0.022***	(0.005)	0.015***	(0.004)	3,801

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

Note: Each regression also includes controls. Control variables are preschool quality, mother's education in years, and household wealth (*z*-score).

Table 7. Gender gaps in EDI explained by difference in education enrollment and parenting (extended decomposition)

	Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Score		Sample Size
		Est	SE	Est	SE	Est	SE	Est	SE	
Physical health & wellbeing	6	0.035	(0.054)	-0.002	(0.008)	0.008	(0.010)	0.030	(0.032)	1,321
	7	0.043	(0.034)	0.000	(0.004)	0.000	(0.004)	0.016	(0.011)	2,860
	8	0.083***	(0.030)	0.002	(0.003)	0.007*	(0.004)	0.023	(0.016)	3,801
Social competence	6	0.155***	(0.054)	0.011	(0.008)	0.004	(0.010)	-0.024	(0.051)	1,321
	7	0.090**	(0.037)	0.019***	(0.005)	0.007	(0.005)	-0.001	(0.030)	2,860
	8	0.260***	(0.032)	0.004	(0.002)	0.016**	(0.007)	0.021	(0.032)	3,801
Emotional maturity	6	0.272***	(0.052)	0.001	(0.007)	0.000	(0.010)	-0.011	(0.043)	1,321
	7	0.228***	(0.036)	-0.007	(0.006)	0.001	(0.006)	0.028	(0.018)	2,860
	8	0.306***	(0.031)	0.002	(0.003)	0.009**	(0.004)	0.036**	(0.017)	3,801
Language & cognitive skills	6	0.161***	(0.037)	0.015*	(0.009)	0.004	(0.010)	-0.018	(0.025)	1,321
	7	0.082***	(0.019)	0.016***	(0.004)	0.006	(0.008)	0.000	(0.006)	2,859
	8	0.089***	(0.011)	0.004**	(0.002)	0.018***	(0.007)	0.007	(0.006)	3,801
Communication skills & general knowledge	6	0.058	(0.056)	0.017	(0.011)	-0.003	(0.008)	0.016	(0.041)	1,321
	7	0.045	(0.039)	0.021***	(0.006)	0.004	(0.006)	-0.016	(0.025)	2,860
	8	0.111***	(0.033)	0.006	(0.004)	0.021***	(0.006)	0.013	(0.015)	3,801

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

Note: Each regression also includes controls. Control variables are center quality, mother's education in years, and household wealth (*z*-score). The extended decomposition model also interacts each control variable with each decision variable (cumulative months of enrollment in preschool, cumulative months in primary, and total parenting score.

Table 8. Gender gaps in EDI explained by difference in education enrollment and parenting by preschool quality (standard decomposition)

		Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Scores	
			Est	SE	Est	SE	Est	SE	Est	SE
High quality preschool	Physical health & wellbeing	6	0.141*	(0.077)	0.010	(0.012)	-0.007	(0.008)	-0.005	(0.012)
		7	-0.050	(0.052)	0.002	(0.007)	0.001	(0.002)	0.022**	(0.010)
		8	0.026	(0.042)	0.008	(0.005)	0.009*	(0.005)	0.019***	(0.007)
	Social competence	6	0.125	(0.093)	0.035**	(0.017)	-0.007	(0.007)	-0.012	(0.032)
		7	0.166***	(0.063)	0.016*	(0.010)	0.009	(0.009)	0.052**	(0.023)
		8	0.228***	(0.054)	0.006	(0.004)	0.005	(0.004)	0.069***	(0.020)
	Emotional maturity	6	0.346***	(0.093)	-0.004	(0.013)	-0.013	(0.012)	-0.010	(0.025)
		7	0.205***	(0.063)	-0.030***	(0.012)	0.002	(0.003)	0.048**	(0.022)
		8	0.320***	(0.052)	-0.000	(0.003)	0.008	(0.005)	0.063***	(0.018)
	Language & cognitive skills	6	0.169***	(0.059)	0.042**	(0.017)	-0.005	(0.005)	-0.003	(0.008)
		7	0.114***	(0.029)	0.020***	(0.007)	0.009	(0.009)	0.004	(0.003)
		8	0.057***	(0.016)	0.007*	(0.004)	0.014**	(0.005)	0.004**	(0.002)
	Communication skills & general knowledge	6	0.027	(0.097)	0.055**	(0.024)	-0.000	(0.005)	-0.004	(0.011)
		7	0.062	(0.065)	0.035***	(0.013)	0.010	(0.010)	0.015*	(0.008)
		8	0.115**	(0.054)	0.013*	(0.008)	0.018**	(0.008)	0.022***	(0.008)
Medium quality preschool	Physical health & wellbeing	6	-0.073	(0.090)	-0.001	(0.005)	0.006	(0.007)	0.005	(0.013)
		7	0.135**	(0.055)	-0.003	(0.006)	0.005	(0.004)	0.002	(0.010)
		8	0.093*	(0.052)	0.002	(0.003)	0.004	(0.005)	0.009	(0.008)
	Social competence	6	0.194**	(0.091)	0.000	(0.002)	0.001	(0.005)	0.010	(0.025)
		7	0.004	(0.065)	0.024**	(0.010)	0.014	(0.009)	0.002	(0.013)
		8	0.299***	(0.052)	0.002	(0.003)	0.017**	(0.007)	0.018	(0.015)
	Emotional maturity	6	0.154*	(0.085)	-0.000	(0.001)	0.003	(0.005)	0.011	(0.029)
		7	0.198***	(0.060)	0.002	(0.007)	0.003	(0.004)	0.003	(0.016)
		8	0.365***	(0.050)	0.005	(0.004)	0.013**	(0.006)	0.017	(0.013)
	Language & cognitive skills	6	0.096*	(0.057)	0.002	(0.009)	0.008	(0.008)	0.004	(0.010)
		7	0.056*	(0.032)	0.020***	(0.007)	0.018*	(0.010)	0.001	(0.003)
		8	0.073***	(0.017)	0.003	(0.002)	0.014***	(0.005)	0.002	(0.001)
	Communication skills & general knowledge	6	0.059	(0.098)	0.002	(0.011)	0.004	(0.006)	0.004	(0.011)
		7	0.060	(0.071)	0.030**	(0.012)	0.012	(0.008)	0.001	(0.006)
		8	0.088	(0.057)	0.007	(0.006)	0.026**	(0.010)	0.009	(0.007)
Low quality preschool	Physical health & wellbeing	6	0.042	(0.107)	-0.011	(0.013)	0.005	(0.007)	0.040*	(0.021)
		7	0.053	(0.066)	0.003	(0.004)	-0.000	(0.000)	0.025**	(0.011)
		8	0.159***	(0.061)	0.000	(0.001)	0.005	(0.005)	0.026**	(0.012)
	Social competence	6	0.143	(0.096)	0.027*	(0.016)	0.001	(0.004)	0.049**	(0.023)
		7	0.093	(0.063)	0.010	(0.008)	-0.001	(0.007)	0.028**	(0.012)
		8	0.254***	(0.059)	0.003	(0.005)	0.021**	(0.010)	0.024**	(0.011)
	Emotional maturity	6	0.319***	(0.090)	-0.007	(0.010)	0.000	(0.003)	0.070**	(0.031)
		7	0.282***	(0.061)	0.001	(0.003)	0.000	(0.001)	0.046**	(0.019)
		8	0.242***	(0.056)	0.001	(0.003)	0.003	(0.004)	0.047**	(0.020)
	Language & cognitive skills	6	0.225***	(0.073)	0.024*	(0.013)	0.008	(0.011)	0.024*	(0.013)
		7	0.079**	(0.035)	0.009	(0.006)	-0.001	(0.010)	0.002	(0.003)
		8	0.146***	(0.023)	0.001	(0.002)	0.016**	(0.008)	0.004*	(0.002)
	Communication skills & general knowledge	6	0.089	(0.098)	0.006	(0.011)	-0.001	(0.004)	0.011	(0.012)
		7	0.017	(0.068)	0.005	(0.005)	-0.001	(0.007)	0.005	(0.006)
		8	0.135**	(0.062)	0.001	(0.002)	0.022**	(0.011)	0.012*	(0.007)

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

Note: Each regression also includes controls. Control variables include mother's education and household wealth (*z*-score).

Table 9. Gender gaps in EDI explained by difference in education enrollment and parenting by preschool quality (extended decomposition)

		Age	Gender difference		Cum. months of enrollment in preschool		Cum. months of enrollment in primary		Total Parenting Scores	
			Est	SE	Est	SE	Est	SE	Est	SE
High quality preschool	Physical health & wellbeing	6	0.141*	(0.077)	0.008	(0.019)	0.005	(0.011)	-0.006	(0.016)
		7	-0.050	(0.052)	-0.005	(0.010)	0.002	(0.005)	0.013	(0.011)
		8	0.026	(0.042)	0.009	(0.010)	0.001	(0.008)	0.018	(0.012)
	Social competence	6	0.125	(0.093)	0.015	(0.023)	-0.012	(0.017)	-0.005	(0.014)
		7	0.166***	(0.063)	0.002	(0.013)	0.011	(0.011)	0.046**	(0.022)
		8	0.228***	(0.054)	0.005	(0.004)	0.022*	(0.013)	0.078***	(0.025)
	Emotional maturity	6	0.346***	(0.093)	0.015	(0.023)	0.006	(0.014)	-0.001	(0.005)
		7	0.205***	(0.063)	-0.035**	(0.014)	0.007	(0.008)	0.030*	(0.017)
		8	0.320***	(0.052)	-0.001	(0.005)	0.016	(0.012)	0.060***	(0.021)
	Language & cognitive skills	6	0.169***	(0.059)	0.010	(0.021)	0.012	(0.014)	-0.005	(0.013)
		7	0.114***	(0.029)	0.009	(0.008)	0.014	(0.013)	0.007	(0.007)
		8	0.057***	(0.016)	0.005	(0.004)	0.023**	(0.010)	0.013**	(0.005)
	Communication skills & general knowledge	6	0.027	(0.097)	0.011	(0.038)	0.008	(0.016)	-0.010	(0.025)
		7	0.062	(0.065)	0.031**	(0.015)	0.013	(0.014)	0.024*	(0.015)
		8	0.115**	(0.054)	0.015	(0.010)	0.020	(0.013)	0.033**	(0.016)
Medium quality preschool	Physical health & wellbeing	6	-0.073	(0.090)	-0.003	(0.013)	0.014	(0.017)	0.006	(0.017)
		7	0.135**	(0.055)	-0.015	(0.010)	0.011	(0.010)	0.002	(0.015)
		8	0.093*	(0.052)	-0.001	(0.005)	-0.002	(0.011)	0.008	(0.008)
	Social competence	6	0.194**	(0.091)	-0.012	(0.023)	0.002	(0.010)	0.010	(0.026)
		7	0.004	(0.065)	0.028**	(0.013)	0.020	(0.014)	0.001	(0.006)
		8	0.299***	(0.052)	0.005	(0.004)	0.008	(0.010)	0.027	(0.022)
	Emotional maturity	6	0.154*	(0.085)	-0.026	(0.020)	0.009	(0.012)	0.010	(0.027)
		7	0.198***	(0.060)	-0.008	(0.013)	0.022	(0.015)	0.002	(0.011)
		8	0.365***	(0.050)	0.007	(0.005)	0.006	(0.010)	0.009	(0.008)
	Language & cognitive skills	6	0.096*	(0.057)	0.001	(0.014)	0.011	(0.012)	0.002	(0.005)
		7	0.056*	(0.032)	0.017**	(0.009)	0.028*	(0.016)	0.001	(0.004)
		8	0.073***	(0.017)	0.005	(0.003)	0.026**	(0.011)	0.003	(0.003)
	Communication skills & general knowledge	6	0.059	(0.098)	0.007	(0.017)	-0.003	(0.010)	0.006	(0.015)
		7	0.060	(0.071)	0.022	(0.018)	0.005	(0.010)	0.000	(0.001)
		8	0.088	(0.057)	0.005	(0.007)	0.023	(0.014)	0.012	(0.010)
Low quality preschool	Physical health & wellbeing	6	0.042	(0.107)	-0.005	(0.016)	0.014	(0.020)	0.055	(0.039)
		7	0.053	(0.066)	0.004	(0.006)	-0.001	(0.009)	0.039*	(0.021)
		8	0.159***	(0.061)	0.002	(0.003)	0.019	(0.013)	0.040**	(0.020)
	Social competence	6	0.143	(0.096)	0.022	(0.016)	0.008	(0.014)	0.008	(0.032)
		7	0.093	(0.063)	0.009	(0.009)	-0.002	(0.014)	0.046**	(0.022)
		8	0.254***	(0.059)	0.002	(0.006)	0.026*	(0.015)	0.031*	(0.017)
	Emotional maturity	6	0.319***	(0.090)	-0.011	(0.014)	0.002	(0.007)	0.061	(0.038)
		7	0.282***	(0.061)	-0.000	(0.009)	-0.001	(0.009)	0.059**	(0.027)
		8	0.242***	(0.056)	0.002	(0.005)	0.014	(0.011)	0.053**	(0.024)
	Language & cognitive skills	6	0.225***	(0.073)	0.020	(0.015)	0.015	(0.020)	-0.025	(0.026)
		7	0.079**	(0.035)	0.011	(0.008)	-0.002	(0.015)	0.002	(0.008)
		8	0.146***	(0.023)	0.001	(0.002)	0.023*	(0.012)	0.007	(0.005)
	Communication skills & general knowledge	6	0.089	(0.098)	0.005	(0.013)	0.009	(0.014)	-0.008	(0.029)
		7	0.017	(0.068)	0.004	(0.007)	-0.002	(0.015)	0.033*	(0.019)
		8	0.135**	(0.062)	0.004	(0.005)	0.040**	(0.020)	0.032*	(0.017)

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

Note: Each regression also includes controls. Control variables include mother's education and household wealth (*z*-score). The extended decomposition model also interacts each control variable with each decision variable (cumulative months of enrollment in preschool, cumulative months in primary, and total parenting score).

Appendix

A1. Regression results for test score on gender (at each age)

			Girl = 1		Constant		N	R-sq
DV: Language (std)	(5)	at age 6	0.252***	(0.052)	-0.089**	(0.037)	1428	0.016
	(6)	at age 7	0.201***	(0.036)	0.328***	(0.026)	3109	0.010
	(7)	at age 8	0.158***	(0.027)	0.803***	(0.019)	4306	0.008
	(8)	at age 9	0.160***	(0.037)	0.988***	(0.026)	2015	0.009
DV: Math (std)	(9)	at age 6	0.243***	(0.052)	-0.071*	(0.038)	1428	0.015
	(10)	at age 7	0.155***	(0.035)	0.373***	(0.025)	3109	0.006
	(11)	at age 8	0.176***	(0.027)	0.623***	(0.020)	4306	0.010
	(12)	at age 9	0.168***	(0.036)	0.810***	(0.025)	2015	0.011
DV: Abstract reasoning (std)	(13)	at age 6	0.107**	(0.053)	-0.014	(0.039)	1428	0.003
	(14)	at age 7	-0.061	(0.039)	0.383***	(0.028)	3109	0.001
	(15)	at age 8	0.003	(0.032)	0.302***	(0.023)	4306	0.000
	(16)	at age 9	0.036	(0.049)	0.435***	(0.035)	2015	0.000

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

Note: Standard errors in parentheses. Each row is the result of separate regressions.

A2. Regression results for EDI scores on gender (at each age)

			Girl = 1		Constant		N	R-sq
DV: Physical health & well-being (std)	(1)	at age 6	0.032	(0.051)	0.013	(0.037)	1409	0.000
	(2)	at age 7	0.031	(0.033)	0.112***	(0.023)	3091	0.000
	(3)	at age 8	0.068**	(0.028)	0.188***	(0.021)	4153	0.001
DV: Social competence (std)	(4)	at age 6	0.166***	(0.052)	-0.415***	(0.037)	1409	0.007
	(5)	at age 7	0.086**	(0.035)	-0.250***	(0.025)	3091	0.002
	(6)	at age 8	0.247***	(0.031)	-0.169***	(0.022)	4153	0.016
DV: Emotional maturity (std)	(7)	at age 6	0.279***	(0.050)	-0.184***	(0.037)	1409	0.022
	(8)	at age 7	0.213***	(0.034)	-0.092***	(0.025)	3091	0.012
	(9)	at age 8	0.296***	(0.029)	-0.076***	(0.021)	4153	0.024
DV: Language & cognitive skills (std)	(10)	at age 6	0.157***	(0.036)	0.288***	(0.027)	1409	0.013
	(11)	at age 7	0.077***	(0.018)	0.553***	(0.013)	3090	0.006
	(12)	at age 8	0.089***	(0.010)	0.710***	(0.009)	4153	0.017
DV: Communication skills & general knowledge (std)	(13)	at age 6	0.070	(0.054)	-0.694***	(0.039)	1409	0.001
	(14)	at age 7	0.056	(0.038)	-0.558***	(0.026)	3091	0.001
	(15)	at age 8	0.094***	(0.032)	-0.267***	(0.023)	4153	0.002

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

Note: Standard errors in parentheses. Each row is the result of separate regressions.

A.3 Regression results for rate and months of enrollment in preschool and primary school on gender (at each age)

			Girl = 1		Constant		N	R-sq
Ever enrolled in preschool (Yes=1)?	(1)	at age 2	0.016	(0.011)	0.036***	(0.007)	1428	0.002
	(2)	at age 3	0.040***	(0.010)	0.121***	(0.007)	4537	0.003
	(3)	at age 4	0.067***	(0.010)	0.337***	(0.007)	8843	0.005
	(4)	at age 5	0.057***	(0.009)	0.613***	(0.007)	10858	0.004
	(5)	at age 6	-0.016*	(0.009)	0.370***	(0.007)	10858	0.000
	(6)	at age 7	-0.008**	(0.003)	0.029***	(0.002)	9430	0.001
	(7)	at age 8	0.000	(0.001)	0.001*	(0.001)	6321	0.000
	(8)	at age 9			-		2015	.
Months enrolled in preschool	(9)	at age 2	0.139	(0.103)	0.346***	(0.069)	1428	0.001
	(10)	at age 3	0.369***	(0.101)	1.170***	(0.067)	4537	0.003
	(11)	at age 4	0.682***	(0.101)	3.295***	(0.070)	8843	0.005
	(12)	at age 5	0.582***	(0.092)	6.063***	(0.066)	10858	0.004
	(13)	at age 6	-0.147	(0.092)	3.674***	(0.065)	10858	0.000
	(14)	at age 7	-0.083***	(0.032)	0.289***	(0.024)	9430	0.001
	(15)	at age 8	0.002	(0.007)	0.007*	(0.004)	6321	0.000
	(16)	at age 9			-		2015	.
Ever enrolled in primary (Yes=1)?	(17)	at age 2			-		1428	.
	(18)	at age 3			-		4537	.
	(19)	at age 4	-0.000	(0.001)	0.001**	(0.001)	8843	0.000
	(20)	at age 5	0.002	(0.003)	0.031***	(0.002)	10858	0.000
	(21)	at age 6	0.054***	(0.010)	0.504***	(0.007)	10858	0.003
	(22)	at age 7	0.019***	(0.004)	0.947***	(0.003)	9430	0.002
	(23)	at age 8	0.002	(0.002)	0.995***	(0.001)	6321	0.000
	(24)	at age 9			-		2015	0.000
Months enrolled in primary	(25)	at age 2			-		1428	.
	(26)	at age 3			-		4537	.
	(27)	at age 4	-0.004	(0.007)	0.013**	(0.005)	8843	0.000
	(28)	at age 5	0.017	(0.033)	0.305***	(0.023)	10858	0.000
	(29)	at age 6	0.521***	(0.091)	4.662***	(0.064)	10858	0.003
	(30)	at age 7	0.176***	(0.048)	8.565***	(0.036)	9430	0.001
	(31)	at age 8	0.044	(0.040)	8.097***	(0.029)	6321	0.000
	(32)	at age 9			-		2015	.

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

Note: Standard errors in parentheses. Each row is the result of separate regressions. For enrollment in preschool, outcome is perfectly predicted at age 9 since none are enrolled. For enrollment in primary, outcomes is perfectly predicted at ages 2 and 3 since none are enrolled, and at age 9 since all are enrolled.

A.4 Regression results for rate and months of enrollment in preschool on gender (at each age) by quality of preschool

				Girl = 1		Constant		N	R-sq
DV: Ever enrolled in preschool (Yes=1)?	High quality preschool	(1)	at age 2	0.040	(0.025)	0.061***	(0.016)	475	0.005
		(2)	at age 3	0.072***	(0.021)	0.164***	(0.014)	1498	0.008
		(3)	at age 4	0.096***	(0.018)	0.394***	(0.012)	3054	0.009
		(4)	at age 5	0.066***	(0.015)	0.656***	(0.011)	3823	0.005
		(5)	at age 6	-0.018	(0.016)	0.420***	(0.011)	3823	0.000
		(6)	at age 7	-0.000	(0.006)	0.029***	(0.004)	3348	0.000
		(7)	at age 8	0.000	(0.002)	0.002	(0.001)	2325	0.000
	Medium quality preschool	(8)	at age 2	0.007	(0.014)	0.021**	(0.009)	481	0.001
		(9)	at age 3	0.037**	(0.017)	0.098***	(0.011)	1501	0.003
		(10)	at age 4	0.057***	(0.017)	0.308***	(0.012)	2965	0.004
		(11)	at age 5	0.075***	(0.016)	0.590***	(0.011)	3609	0.006
		(12)	at age 6	-0.030*	(0.016)	0.363***	(0.011)	3609	0.001
		(13)	at age 7	-0.010*	(0.006)	0.031***	(0.004)	3128	0.001
		(14)	at age 8	0.001	(0.001)	0.000***	(0.000)	2108	0.000
	Low quality preschool	(15)	at age 2	0.002	(0.015)	0.027**	(0.011)	472	0.000
		(16)	at age 3	0.012	(0.016)	0.101***	(0.011)	1538	0.000
		(17)	at age 4	0.049***	(0.018)	0.306***	(0.012)	2824	0.003
		(18)	at age 5	0.031*	(0.017)	0.589***	(0.012)	3426	0.001
		(19)	at age 6	0.005	(0.016)	0.321***	(0.011)	3426	0.000
		(20)	at age 7	-0.015***	(0.005)	0.029***	(0.004)	2954	0.003
		(21)	at age 8	-0.000	(0.001)	0.001	(0.001)	1888	0.000
DV: Months enrolled in preschool	High quality preschool	(22)	at age 2	0.324	(0.234)	0.583***	(0.153)	475	0.004
		(23)	at age 3	0.670***	(0.200)	1.578***	(0.132)	1498	0.007
		(24)	at age 4	0.972***	(0.177)	3.833***	(0.122)	3054	0.010
		(25)	at age 5	0.661***	(0.149)	6.508***	(0.108)	3823	0.005
		(26)	at age 6	-0.170	(0.159)	4.182***	(0.112)	3823	0.000
		(27)	at age 7	-0.008	(0.057)	0.282***	(0.040)	3348	0.000
		(28)	at age 8	0.003	(0.014)	0.012	(0.008)	2325	0.000
	Medium quality preschool	(29)	at age 2	0.051	(0.138)	0.213**	(0.094)	481	0.000
		(30)	at age 3	0.324**	(0.162)	0.964***	(0.106)	1501	0.003
		(31)	at age 4	0.599***	(0.172)	3.020***	(0.117)	2965	0.004
		(32)	at age 5	0.764***	(0.160)	5.835***	(0.114)	3609	0.006
		(33)	at age 6	-0.298*	(0.158)	3.604***	(0.112)	3609	0.001
		(34)	at age 7	-0.098*	(0.056)	0.303***	(0.043)	3128	0.001
		(35)	at age 8	0.006	(0.006)	-0.000***	(0.000)	2108	0.000
	Low quality preschool	(36)	at age 2	0.039	(0.146)	0.244**	(0.101)	472	0.000
		(37)	at age 3	0.128	(0.154)	0.973***	(0.107)	1538	0.000
		(38)	at age 4	0.481***	(0.175)	2.993***	(0.122)	2824	0.003
		(39)	at age 5	0.332**	(0.167)	5.797***	(0.119)	3426	0.001
		(40)	at age 6	0.066	(0.159)	3.163***	(0.113)	3426	0.000
		(41)	at age 7	-0.148***	(0.052)	0.283***	(0.043)	2954	0.003
		(42)	at age 8	-0.003	(0.013)	0.011	(0.011)	1888	0.000

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

Note: Standard errors in parentheses. Each row is the result of separate regressions. For enrollment in preschool, outcome is perfectly predicted at age 9 since none are enrolled. For enrollment in primary, outcomes is perfectly predicted at ages 2 and 3 since none are enrolled, and at age 9 since all are enrolled.

A5. Regression results for parenting score on gender (at each age)

			Girl = 1		Constant		N	R-sq
DV: Parenting score (std)	(1)	at age 6	0.076	(0.054)	0.021	(0.039)	1413	0.001
	(2)	at age 7	0.114***	(0.035)	-0.005	(0.024)	3094	0.004
	(3)	at age 8	0.141***	(0.030)	-0.086***	(0.022)	4291	0.005
	(4)	at age 9	0.134***	(0.046)	-0.140***	(0.033)	2007	0.004

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

Note: Standard errors in parentheses. Each row is the result of separate regressions.