Parental investment, child's gender and skill gaps [1]

Anastasiia Suvorova [2]

August 21, 2024

Latest version available here.

Abstract

Female-favourable gender gaps in multiple measures of academic achievement, among children and young adults, have increased over recent years. These disparities have been linked to deficits in boys' literacy, which manifest early in childhood and accumulate over time. I investigate the impact of parental time investment decisions on the widening of literacy gaps between boys and girls. I estimate a model of the mother's time investment and child skill accumulation, allowing for gender differences in the literacy production function, initial endowments, and parental preferences for children's human capital. My analysis centers on the development of skills in children aged 6 to 15 as observed in the Longitudinal Study of Australian Children. I document that mothers tend to allocate more time to their daughters, and although my point estimates suggest that these differences in maternal time investment are mainly explained by the mother's preferences for literacy between boys and girls, these estimates are not statistically significant. Overall, my findings suggest that the role of mothers' aggregate time investments in the expansion of literacy deficits may be limited, and this expansion is driven by productivity differences unrelated to aggregate time investment.

[1] I would like to express my gratitude to Audra Bowlus, Lance Lochner, and Sergio Ocampo Diaz, for their guidance and support. I am thankful to Rory McGee, Nirav Mehta, and participants of the AMM Reading Group for their excellent advice. All errors are my own.

[2] Department of Economics, Western University, London, ON, N6A 5C2, e-mail: asuvorov@uwo.ca

1 Introduction

The successful accumulation of literacy during childhood is an important factor in determining the educational outcomes of young adults. The empirical literature on childhood development has documented a female-favorable gender gap in literacy and language development occurring before kindergarten and expanding through the school years (Cobb-Clark and Moschion, 2017). These deficits are associated with female-favorable gender gaps in several measures of academic achievement, including the probabilities of being retained in grade, graduating from high school, and graduating from college (Buchmann and DiPrete, 2006, Entwisle, Alexander, and Olson, 2007, Autor, Figlio, Karbownik, Roth, and Wasserman, 2020).

Literature in multiple disciplines including economics and development psychology recognizes the critical role of mothers, particularly mother's time investment in various learning-conducive activities, in nurturing skill accumulation in children. The applied structural literature in economics has developed multiple critical insights into the determinants of a mother's time investment and the role these decisions have in the inequality of child development trajectories. However, it mainly abstracts from inequality across children's gender. This paper leverages a structural model of household production and child development, and estimates it using the Longitudinal Survey of Australian Children (LSAC) data to quantify the contribution of various factors to the expansion of female-favourable gaps in literacy in school-aged children. My findings show that the role of the mother's aggregate time in explaining the expansion of literacy gaps in children throughout the school age is limited.

Several explanations have been proposed to explain the apparent deficits in boys' academic achievement. Human capital during childhood is shaped by multiple properties of early life environment interacting with genetic endowments (Almond, Currie, and Duque, 2018). These genetic endowments can differ across boys and girls and affect children's initial ability for literacy and communication. For example, more areas of girls' brains are dedicated to verbal functioning and girls tend to have stronger neural connectors in their temporary lobes which facilitates listening (Michael and Stevens, 2010). Because skills are self-productive, having a higher endowment of literacy by the start of formal schooling can allow girls to learn at a faster rate throughout the school years (Cunha and Heckman, 2008).

It is also possible that boys and girls learn in different ways and, therefore, vary in their production functions of skills. These differences in learning can have a physiological nature. For example, boys tend to have more dopamine in their bloodstream, which limits their ability to learn while being sedentary (Michael and Stevens, 2010). Children can also learn differently due to societal values and norms. For example, girls' advantage in reading and math is positively correlated with measures of gender equality in their countries of origin (Guiso, Monte, Sapienza, and Zingales, 2008, Rodríguez-Planas and Nollenberger, 2018). Moreover, gender-based stereotypes by teachers and parents can influence learning, long-term student outcomes and choices (Nicoletti, Sevilla, and Tonei, 2022, Lavy and Sand, 2015).

Alternatively, female-favourable gaps in literacy can emerge due to the gaps in learning environments that can propagate differential developmental trends. Specifically, parental time investment is commonly recognized as an important input in the production of children's cognitive skills (Del Boca, Flinn, and Wiswall, 2014, Fiorini and Keane, 2014, Caucutt, Lochner, Mullins, and Park, 2020), and an increasing amount of evidence suggests that parents allocate their time and resources differently depending on the child's gender. For example, Baker and Milligan (2016) document that in Canada, the United States, and the United Kingdom parents of children aged zero to three read more often with their girls, which explains as much as a third of the gender gap in reading scores in the early grades. Nicoletti and Tonei (2020) show evidence of differences in time investment strategies adopted for daughters and sons, with parents compensating more for negative shocks to the cognitive skills of daughters compared to negative shocks to the cognitive skills of sons. Bertrand and Pan (2013) show that parents are significantly more likely to read to their girls, take their girls to a concert, and to sign them up for an extracurricular activity.

While this expanding evidence suggests that parents tend to consistently treat their sons and daughters differently, it still remains to be understood why parents make different time investment choices for their sons and daughters. The structural literature exploring family investment decisions and child skill accumulation proposes three main determinants of parental time investment: preferences for the human capital of their children or investment itself, budget and time constraints, and production function of child human capital. While recognizing the size of the gender gaps in skills and parental investments, literature exploring gender gaps does not conclude whether these gaps are driven by differences in the skill production process for boys and girls, or by gendered parental preferences for the human capital (Lundberg, 2017). This paper quantifies the contribution of gender-specific differences in the production function of literacy, preferences, and initial endowments in explaining the gaps in mother's time investment and expanding female-favourable gaps in literacy. To do so I augment the model of Del Boca, Flinn, and Wiswall (2014) by allowing the heterogeneity across a child's gender, and estimate it using the data for school-aged children in the K-cohort of the Longitudinal Survey of Australian Children (LSAC).

I document that mothers tend to allocate more time to their daughters. My point estimates suggest that boys benefit more from the mother's time investment, which implies that differences in maternal time investment are mainly explained by the mother's preferences for the literacy of girls or for the time spent with their daughters. However, these estimates are not statistically significant. Overall, my findings do not suggest that inequality in the mother's aggregate time investment plays an important role in the expansion of literacy deficits in school-aged children. By contrast, the growth of the gender gap is driven by other productivity differences between boys and girls.

This paper contributes to the structural literature exploring the sources of inequality in parental investment decisions across different groups of children characterized by race, family income, mother's education and socioeconomic status.[3] My paper focuses on the heterogeneity of parental investment decisions with respect to the child's gender. Gugl and Welling (2012) develop and solve a theoretical model of parental time investment in sons and daughters, however, they do not estimate their model to quantify the contribution of various channels. My paper fills this gap by using structural modelling to understand differential parental investment strategies across

[3]For example, Cunha (2014) quantifies the role of differences in preferences, subjective beliefs, budget constraints and marginal returns to investment in explaining the racial gaps in early investment in children. Caucutt and Lochner (2020) focus on the contribution of borrowing constraints and the intergenerational transmission of ability in explaining the inequality of parental investment and child achievement across families with different income levels. Brilli (2022) explores the role of differences in the productivity of time investment across mothers with different education levels.

sons and daughters.

I also contribute to the expanding literature exploring statistical analysis to shed light on the source of gender differences in the academic outcomes of boys and girls. Important insights from this literature include a higher sensitivity of the development of boys to family investment and family advantage (Autor, Figlio, Karbownik, Roth, and Wasserman, 2020, Chetty, Hendren, Lin, Majerovitz, and Scuderi, 2016, Brenøe and Lundberg, 2018) as well as documented gender-related differences in various family investment decisions (Baker and Milligan, 2016, Nicoletti and Tonei, 2020, Bertrand and Pan, 2013). In contrast to this literature, I take a structural approach to quantify the drivers behind inequality in the family environment.

This paper proceeds as follows. Section 2 provides details on a dynamic model which predicts how mothers' time investment and child development depends on the production function of skills, parental preferences for human capital, budget constraints and endowments. Sections 3 and 4 describe data and how it is used to estimate the parameters of the model. Section 5 discusses the implications of my estimates for understanding the gaps in mother's time investment and literacy across boys and girls, and Section 6 concludes.

2 Model

To quantify the sources of expanding gender gaps in literacy, I build on the model of human capital accumulation and household choices developed in Del Boca, Flinn, and Wiswall (2014), and augment it to allow for gender-specific heterogeneity in the production function, initial endowments, and parental preferences for human capital. In the model, mothers are the only decision-makers about investment in children, and they are deciding on the optimal investment in one child with the total fertility exogenous. The investments and skills in the survey are measured only for one child per household, therefore, the gender of the child in the model corresponds to the gender of the only child in the household for whom investments and skills are available.

2.1 Mother's preferences

The model begins when the child is 6 to 7 years old, and receives an initial draw of human capital from a gender-specific distribution $S_1^j \sim N(\mu_1^j, \sigma_1^j)$ for boys and girls $j = b,g$. There are 5 periods in the model $M = 5$, and every period corresponds to a two-year age category of children. From the first period $t = 1$ corresponding to children ages 6-7 to the last period $t = M$ corresponding to children ages 14-15 mothers choose hours of work (h_t^m) and hours of active time invested in children (τ_t^m). In period $t = M + 1$ the child begins a new stage of development.

Mother's period utility depends on hours of leisure (l_t^m), consumption (c_t), and child's human capital (S_t):

$$U(l_t^m, c_t, S_t) = \alpha^{j,m} \ln l_t^m + \alpha^{j,s} \ln S_t + \alpha_t^{j,c} \ln(c_t), \tag{1}$$

where $\alpha^{j,s}$, $\alpha^{j,c}$, $\alpha^{j,m}$ reflect preferences for the child's human capital, consumption, and leisure respectively, and $\alpha^{j,s}+\alpha^{j,m}+\alpha^{j,c} = 1$. I introduce heterogeneity with respect to the child's gender into the utility function by allowing the preference parameters to be different across mothers of boys and girls $j = \{b,g\}$.

2.2 Human capital production function

The production function of literacy captures the complementarity and self-productivity of human capital investment. Human capital in $t + 1$ is produced by the active time investment by the mother and human capital in period t. It evolves according to a Cobb-Douglas production technology

$$\ln S_{t+1} = \ln R_t^j + \rho_t^{j,m} \ln \tau_t^m + \rho_t^{j,s} \ln S_t, \tag{2}$$

where R_t^j is the total factor productivity parameter, $\rho_t^{j,m}$ is age-specific active time productivity, and $\rho_t^{j,s}$ is the human capital self-productivity parameter.

I assume that productivity parameters change monotonically with age (Del Boca, Flinn, and Wiswall, 2014).[4] The heterogeneity with respect to gender is introduced by allowing boys

[4]Since I only observe investment in one child, I have estimated a version of the model allowing the productivity of inputs to differ depending on the number of siblings. For example, it is possible that in families with multiple

and girls to have different production function parameters. Input-specific productivities can have different slope and age dynamics for boys and girls. These are given by

$$R_t^j = exp(\gamma^j + \gamma^j_{age} t + \nu^j_{it}), \tag{3}$$

where $\nu^j_{it} \sim i.i.d.(0,\, \sigma^{j,\nu})$ is a development-specific shock, the distribution of which can also differ across boys and girls. The input-specific productivities are restricted to be positive for the time investment ($\rho_t^{j,m}$) and the lagged value of skills ($\rho_t^{j,s}$), such that for $k = m, s$

$$\rho_t^{j,k} = exp(\gamma^{j,k} + \gamma^{j,k}_{age} t). \tag{4}$$

2.3 Dynamic problem

Mothers maximize utility given a vector of state variables $J_t = (S_t, w_t^m, NI_t)$ which includes the period t wage of the mother, the available period t non-labour income of the mother, NI_t, which includes the father's income in two-parent households, and the human capital stock of the child, S_t. The consumption level of the mother is determined according to the period budget constraint which depends on their labour and non-labour income. The allocation of time between leisure, investment in children, and work is determined by the time constraint, with the total time endowment depending on the total number of children in a given period $nsib_t$.

$$\begin{aligned} V_t(J_t) = \max_{h_t^m, \tau_t^m} \quad & U(l_t^m, c_t, S_t) + \beta E_t V_{t+1}(J_{t+1}) \\ \text{s.t.} \quad & c_t = w_t^m h_t^m + NI_t^* \\ & \bar{T} - \theta^j_{sib} \times nsib_t = l_t^m + h_t^m + \tau_t^m, \end{aligned} \tag{5}$$

where mothers have access to a share b^j of the total non-labour income $NI_t^* = b^j_{NI} NI_t$.[5] Here $\beta \in (0,1)$ is a time discount factor, θ^j_{sib} is a time penalty for having children other than the study child, and E_t is the expectation conditional on the information set t. The terminal value function

children, mothers delegate more childrearing tasks to their daughters and mother's time is less productive for girls. In Appendix F I show that the number of siblings does not significantly affect the inputs' productivity.

[5]The total non-labour income includes mother's non-labour income in all families and father's total income in two-parent households.

in the final period M depends on the mother's valuation of human capital [6]

$$V_M = \alpha^{j,s} ln(S_{M+1}). \tag{7}$$

Following Del Boca, Flinn, and Wiswall (2014) and Mullins (2019) I solve the dynamic problem recursively to find the analytical solution for optimal time investment by mothers (τ_t^m) conditional on their labour supply choice (h_t^m):

$$\tau_t^m = (T - h^m)\frac{\phi_t^{j,m}}{\alpha^{j,m} + \phi_t^{j,m}}, \tag{8}$$

where $\phi_t^{j,m} = \beta\rho_t^{j,m}\frac{\partial V_{t+1}}{\partial lnS_{t+1}}$ and $\frac{\partial V_{t+1}}{\partial lnS_{t+1}}$ is found recursively as

$$\frac{\partial V_{M+1}}{\partial lnS_{M+1}} = \alpha^{j,s}, \quad \ldots, \quad \frac{\partial V_t}{\partial lnS_t} = \alpha^{j,s} + \beta\rho_t^{j,s}\frac{\partial V_{t+1}}{\partial lnS_{t+1}}. \tag{9}$$

The model predicts an interior solution for the optimal choice of time investment. Due to the positive self-productivity of skills $\rho_t^{j,s}$, the marginal utility of child quality $\frac{\partial V_t}{\partial lnS_t}$ reflects both the marginal utility of present child quality, $\alpha^{j,s}$, and the discounted value of future utility.[7]

The structure of the model also allows for the analytical solution for the mother's labour supply. The mother participates in the labour market depending on the latent labour supply level $\tilde{h}_t^m$:

$$h_t^m = \begin{cases} 0 & \text{if } \tilde{h}_t^m \leq 0 \\ \tilde{h}_t^m & \text{if } \tilde{h}_t^m > 0 \end{cases}. \tag{10}$$

The latent labour supply, in turn, depends on the mother's wage, the available non-labour income and a combination of parameters related to mothers' preferences and the production function of skills

[6] I have estimated a more flexible version of the model where the terminal value function in the final period M depends on the mother's valuation of the final human capital Ω^j

$$V_M = \frac{\alpha^{j,s} ln(S_{M+1})}{1 - \Omega^j}. \tag{6}$$

However, Ω^j is very imprecisely estimated with the available data.

[7] See Del Boca, Flinn, and Wiswall (2014) for a detailed description of results.

$$\tilde{h}_t^m = \frac{(T - \theta_{sib}^j \times nsib_t) w_t^m \alpha^{j,c} - (\alpha^{j,m} + \phi_t^{j,m}) NI_t^*}{w_t^m (\alpha^{j,m} + \alpha^{j,c} + \phi_t^{j,m})}. \tag{11}$$

The model predicts that mothers spend less time in paid employment if the return to investing time in children, $\phi_t^{j,m}$ is higher. The return to time investment depends positively on the valuation of human capital, the productivity of time, and the self-productivity of skills. Any of these factors can drive higher returns to parental time investment in girls compared to boys, which might explain the gender differential in time investment. Mother's labour supply is also negatively affected by the preference for leisure, $\alpha^{j,m}$, and the available non-labour income, NI_t^*.

In the presence of increasing evidence that mothers invest more time in their daughters, it is important to quantify what mechanisms of the household choice model drive these diverging investment patterns. Equation (8) shows that there can be multiple drivers playing a role. First, mothers of girls might have a higher valuation of literacy $\alpha^{j,s}$, in which case girls' literacy has a higher marginal utility for the household. Alternatively, parental time investment can be more productive for girls, reflected in $\rho_t^{j,m}$, or literacy can have higher self-productivity for girls, reflected in $\rho_t^{j,s}$, which increases returns to time investment in girls compared to boys. Finally, other factors like the time penalty for having other siblings or the share of available non-labour income can vary across parents of boys and girls if, for example, mothers of boys and girls differ in their bargaining power and access to household income (Pollmann-Schult, 2017). To quantify the role of these various factors in the expansion of the gender gap in literacy I will estimate the model using the data in the Longitudinal Study of Australian Children (LSAC), described in the following section.

3 Data

The data for this project come from the LSAC - a national study of children in Australia that tracks and investigates the effect of childhood environments on children's development and life course trajectories. The LSAC survey measures the evolution of children's literacy, parental time investment and labour choices over more than a decade in children's lives.

The survey commenced in 2004 and participating families were interviewed once every two years. It follows the development of two cohorts of children, the baby cohort (B-cohort) which includes 5107 children aged 0-1 in 2004, and the kindergarten cohort (K-cohort) which includes 4983 children who were aged 4-5 in 2004. I use the K-cohort sample of children due to the availability of consistent measurements of cognitive development and parental time investments when children in the K-cohort are between 6-7 and 14-15 years old.

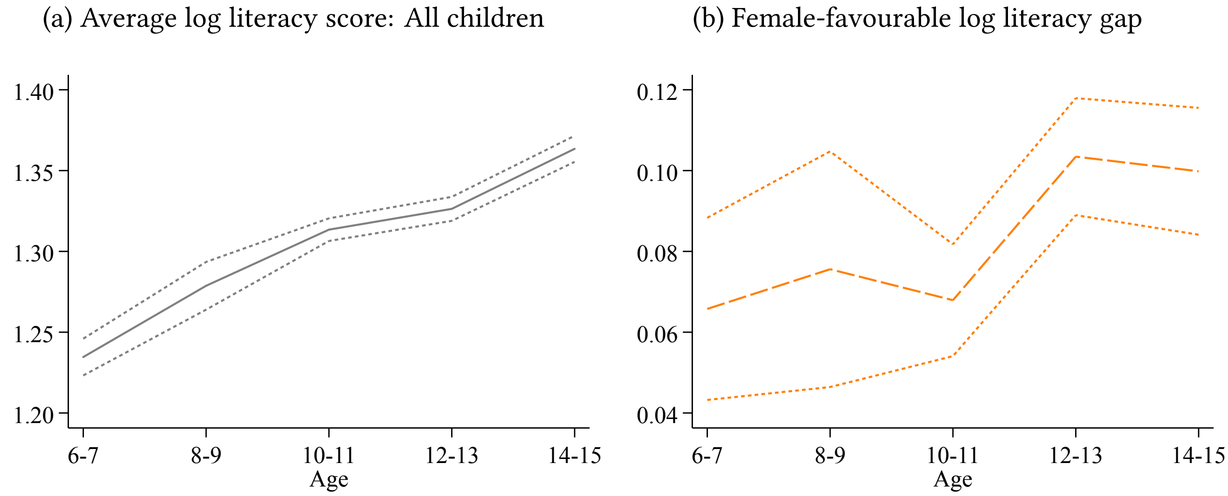
To measure children's literacy, I use the teacher-reported evaluation of literacy available for children of school age: the Academic Rating Scale (ARS). The Academic Rating Scale in Language and Literacy is constructed based on the teacher's answers to nine Language and Literacy items in the questionnaire which evaluate skills, knowledge, and achievement level.[8] The overall measure of achievement is calculated using the Rasch rating score model. The higher scores indicate more advanced levels of literacy. ARS has been shown to validly assess the development of literacy skills in children (Kim and Camilli, 2014). It has been commonly used as a measure of teachers' perceptions about absolute achievement in literacy skills in the previous literature (Cobb-Clark and Moschion, 2017, Robinson and Lubienski, 2011, Cornwell, Mustard, and Van Parys, 2013).

Figure 1 illustrates the trends in the evolution of log literacy score throughout the school age. Panel (a) shows that children accumulate their literacy skills and become increasingly proficient in working with textual information. The literacy score increases by 13% as children age from 6-7 to teenage years. Panel (b) shows that at the start of formal schooling, girls outperform boys in their measures of literacy by 6.5 percent, and this gap expands to 10 percent by 14-15. These expanding female-favourable gaps in literacy were documented for various countries and educational systems (Borgonovi, Choi, and Paccagnella, 2018).

The time investment by mothers is measured based on the time use diary data available for children ages 6 - 15. When children are ages 6-7 and 8-9 the information on time use diaries is collected from one of the caregivers through a self-complete form for two 24-hour periods, one weekday, and one weekend day. Due to the decline in response rates, starting from children

[8]For example, teachers evaluate whether the child conveys ideas when speaking, uses strategies to gain information from print, reads fluently, reads grade-level books, comprehends informational text, composes multi-paragraph texts, redrafts, writes, makes editorial corrections, etc. The teacher is asked to evaluate literacy by choosing from several options: not yet; beginning; in progress; intermediate; proficient.

Figure 1: Average log teacher-reported literacy by children's Age

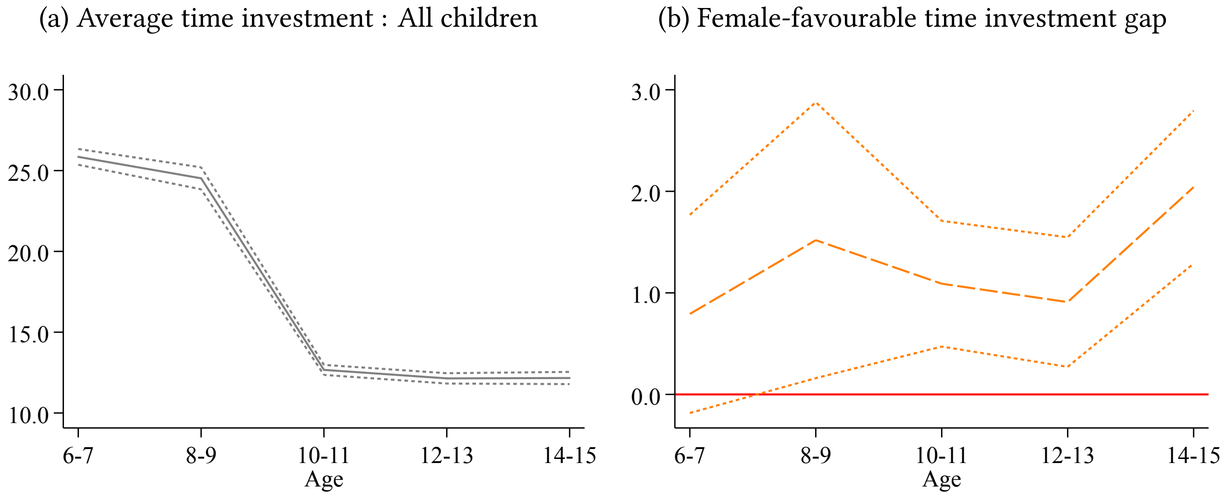


Notes: Panel a) displays the average log teacher-reported literacy score by child's age. Panel b) displays the female-favourable gap in log teacher-reported literacy score by child's age. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income. All panels additionally display 95% confidence intervals.

ages 10-11 the information on time diaries was collected from the child for one day of the week, by asking the study child to record all the activities they have performed during a day and then recovering this information with the assistance of the interviewer. Both time diary formats allow distinguishing between various types of activities the child was engaged in and people who were present with the child.

shorro (see Appendix E). When the time diary information is available for a weekend and a weekday for children ages 6-9, I impute the weekly time investment by multiplying the weekday investment by 5 and the weekend investment by 2 and taking a sum of the two components. When children are older than 10-11 and time diaries are only collected for one day per week, I impute the average weekly time investment using the average share of weekend and weekday time investment in the total weekly time investment estimated for children ages 6-9. On an average weekday, mothers spend around 12.5% of their total weekly active time with children, and on an average weekend day, mothers spend a larger share of around 18.8%.

Figure 2: Mother's weekly active time investment by children's age, hours



Notes: Panel a) displays the average weekly active time investment of mothers by child's age. Panel b) displays the female-favourable gap in weekly active time investment of mothers by child's age. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income. All panels additionally display 95% confidence intervals.

Figure 2 summarizes the trends in hours of weekly active time investment by children's age. Panel (a) shows that the average time investment decreases with age, from over 25 hours per week for children ages 6-7, to around 12 hours per week for teenagers. Panel (b) shows that mothers consistently invest more time in their daughters compared to their sons. On average, mothers invest around 1.23 hours more in engaged activities with their daughters. Given the importance of parental time investment for children's skill accumulation, these gaps in time investment may lead to inequality in skill accumulation.

The survey additionally collects information about a variety of household labour market and demographic characteristics that are important for the understanding of family decisions. When children are ages 6-7, the survey collects only information about the total income without distinguishing between the wage and non-labour income. For children ages 10-11 and above, it is possible to compute hourly wage rates and non-labour income by leveraging the information about total income, salary income, and weekly hours worked. I compute the mother's non-labour

Table 1: Summary of household characteristics by child's age

	Ages 6-7			Ages 8-9			Ages 10-11			Ages 12-13			Ages 14-15		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Mothers hourly wage 2003 dollars				521	22.78	11.63	2192	23.34	12.49	2088	23.75	12.67	1678	24.30	12.64
Non-labour income including spousal wage				698	1183.31	1023.14	3046	1116.82	1172.37	2719	1129.10	1177.46	2105	1126.15	1028.22
Weekly work hours	1410	17.35	16.36	790	20.41	16.00	3513	20.68	16.48	3179	23.56	16.81	2450	25.25	16.92
Mother employed	1410	0.69	0.46	790	0.78	0.42	3508	0.76	0.43	3176	0.80	0.40	2446	0.82	0.38
Mother college+	1410	0.38	0.49	790	0.40	0.49	3513	0.38	0.48	3179	0.39	0.49	2450	0.42	0.49
Mother age	1409	37.55	4.91	789	39.68	5.07	3511	41.13	5.24	3177	43.35	5.04	2449	45.53	5.06
Num. of siblings	1410	1.53	0.96	790	1.61	1.00	3513	1.63	1.06	3179	1.58	1.06	2450	1.51	1.02
Two-parent household	1410	0.87	0.33	790	0.84	0.37	3513	0.79	0.41	3179	0.78	0.41	2450	0.79	0.41

Notes: This table summarizes variables used in the analysis by child's age. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

income by subtracting salary income from the total weekly income and adding the father's weekly total income for households with two parents. I deflate income variables to 2003 Australian dollars and trim the top and bottom 1 percent of hourly wages by the survey wave.

I keep only children ages 6-15 with non-missing measures of the mother's active time investment and labour supply, and children who appeared in the survey at least twice. I additionally limit the sample to mothers with positive active time investment, which represent over 90% of the sample for young children and 80% of the sample for teenagers.[9] For 1.41 percent of reports, the computed available non-labour income of mothers is below zero. If the value of non-labour income is between - 1 and 0, I set it to zero under the assumption that a small difference is due to the discrepancy in retrospective reports. I drop around 1 percent of remaining observations with negative available non-labour income.

Table 1 summarizes the household characteristics of children in the final sample.[10] Around 40 percent of mothers had a college degree, the majority of them were already employed when

[9]The model assumes an interior solution to the optimal time investment, and the production function depends on log weekly active hours, therefore, it requires positive active time investment reports by mothers. See Appendix G for the share of positive active time investment reports by child's age.

[10]See Appendix G for the summary characteristics of a full LSAC sample of children ages 6-15.

children were ages 6-7, and the share of employed mothers increased from 69 to 82 percent by the time children reached 14-15 years old. Due to the availability of spousal incomes for the majority of mothers, they had substantial values of weekly non-labour income averaging over 1000 dollars. Children in the sample were typically not the only children in the household and had between 1 and 2 siblings.

The LSAC dataset contains the measures of the main state variables (parameters of wage and non-labour income processes and child literacy levels) and choice variables (hours of leisure, employment, and time investment in children) that are critical for estimating the model of household investment choice and quantifying the role of these choices in the expansion of gender gaps in literacy. The following section explains how this data is used in estimating the parameters of the model and quantifying the role of various drivers behind the expansion of the gender gap in literacy.

4 Estimation

I estimate the model in two steps. First, I leverage the rich LSAC data to estimate the production function of children's literacy, the mother's wage process, the mother's non-labour income process, and the initial distribution of literacy. Second, given the estimates from the first step, I estimate preference parameters using the Method of Simulated Moments. The standard errors and 90% confidence intervals are computed based on 500 bootstrap repetitions of the estimation process. The following subsections provide details on the model estimation procedure and results.

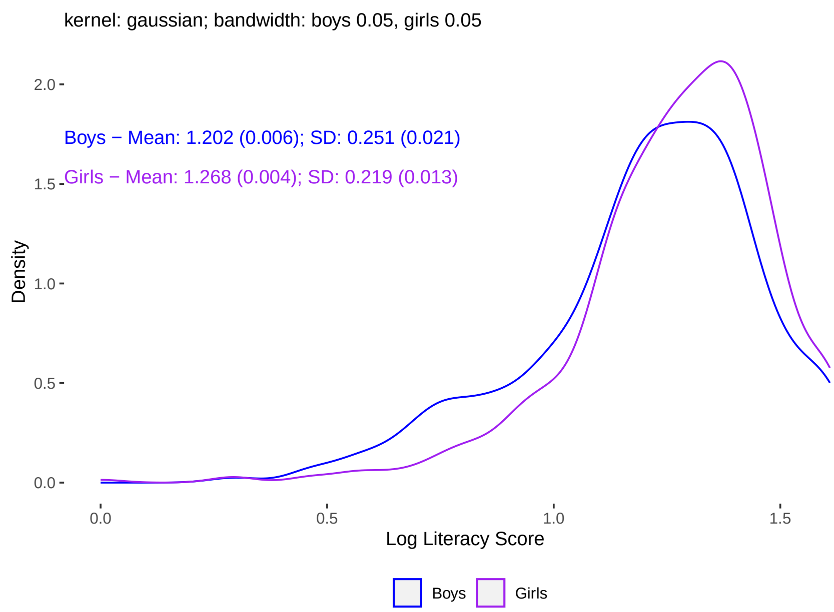
4.1 Initial distribution of literacy and production function

Due to the self-productivity of skills, initial endowments of literacy can play a substantial role in determining how human capital evolves throughout childhood. For example, if a child cannot read words at the beginning of their schooling, while the average student can, they might experience significant delays in learning more advanced literacy skills, such as composing sentences and paragraphs. Moreover, if boys and girls learn differently and young boys, for example, benefit more from family investments and monitoring compared to girls, then not only they will experience higher human capital growth in response to the improvement in the home environment, but it might also be optimal for parents to accommodate these varying learning strategies by adjusting their investment choices. Therefore, it is important to understand when the gender gap in literacy emerges, and how literacy is accumulated by boys and girls. This subsection discusses how the initial distribution of literacy and the production function of literacy are estimated in the data.

The production function of literacy has both inputs (time investment and human capital stock) and outputs (next period value of human capital stock) measured in the data biennially. In the absence of measurement error and unobserved heterogeneity, the production function is identified from the data and can be estimated with a non-linear least squares estimator. Since there is a possibility that the role of all inputs and the age dynamics of the input productivity differs across boys and girls, I estimate production function separately by gender. Parameters of the initial distribution of skill, $S_1 \sim N(\mu_1^j, \sigma_1^j)$, for boys and girls $j = b,\ g$ are also estimated from the data by computing the mean and standard deviation of the literacy score for subsamples of boys and girls ages 6-7.

The model begins when children are ages 6-7 and their teachers first evaluate their literacy. Figure 3 plots the kernel densities for the initial distribution of literacy scores for boys and girls at ages 6-7. Girls arrive at the early ages of formal schooling with higher average levels of log literacy score (the average of 1.268 vs 1.202 for girls and boys respectively). This is consistent with the previous evidence showing that girls have higher averages and lower variability of early literacy levels (Cobb-Clark and Moschion, 2017). Initial literacy levels are also slightly more variable for boys compared to girls, which is consistent with the evidence in Machin and Pekkarinen (2008) for older children. The differences in literacy levels at the start of school can be driven by differences in genetic predisposition to verbal functioning (Michael and Stevens, 2010) or by early childhood and prenatal investments in children or returns to these investments (Almond, Currie, and Duque, 2018), therefore, the role of initial gaps in literacy at ages 6-7 can be attributed to either of these channels. Depending on the extent to which literacy is self-productive, these initial differences in skill endowments can reinforce themselves throughout later learning

Figure 3: Initial distribution of literacy at ages 6-7



Notes: This figure displays the kernel density estimates of the initial distribution of log literacy score for children ages 6-7 by gender along with the estimates of the parameters of the initial distribution. The standard errors are reported in brackets and computed with 500 bootstrap iterations of the estimation procedure. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

in school.

Figure 4 shows the estimates of the production function from Equation (2) obtained with non-linear least squares for subsamples of boys and girls. It illustrates the productivity of various inputs for the next period's literacy levels, computed based on Equations (4) and (3) along with the 90% confidence intervals. Due to small sample sizes the parameters are imprecisely estimated and are mostly not statistically different across boys and girls.[11] Panel (a) shows that the productivity of active time spent with mothers is positive for young boys and declines with the child's age.[12] For girls, the estimated productivity of active time with mothers is slightly smaller with less

[11] The number of child observations in regressions varies across years from just over 120 to over 1000 observations.

[12] Appendix C shows that this dynamics of the productivity of active time is similar if I adjust for the measurement error in log literacy score in estimating production function.

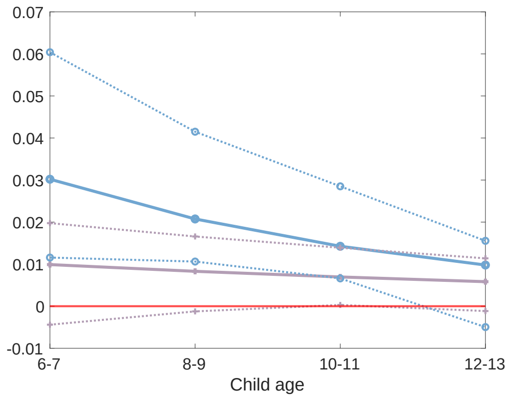
decline over age. This expands on the findings in the previous literature suggesting that a boy's development is more sensitive to the quality of the home environment. For example, Autor, Figlio, Karbownik, Roth, and Wasserman (2020) and Brenøe and Lundberg (2018) show that family disadvantage disproportionately impedes the development of boys during the school age, while Fan, Fang, and Markussen (2015) find that mother's employment during children's early years has a stronger negative effect on the educational outcomes of sons.

The self-productivity of literacy is important for both boys and girls, with slightly higher point estimates for boys. Consistently with these point estimates, Cobb-Clark and Moschion (2017) show that girls score higher on their third-grade reading tests in large part because they were more ready for school at age 4 and had better teacher-assessed literacy skills in kindergarten. While initial endowments commonly play an important role in explaining persistent gaps in literacy, it is not clear whether they can explain the expansion of the gender gap in literacy throughout the school age.

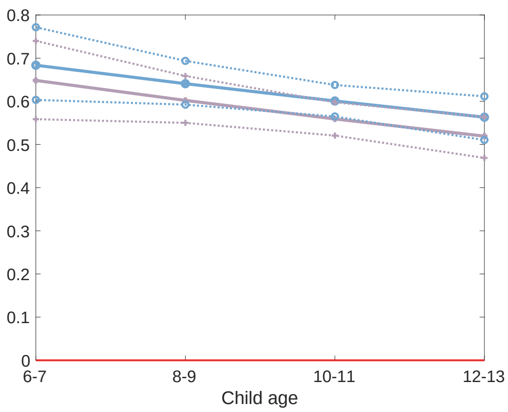
Importantly, the role of total factor productivity is substantial and increasing with age. It is also higher for girls, with the difference remaining stable over the age of children and statistically significant at ages 10-11. This is driven by a substantial part of the growth in girls' literacy scores not explained by mothers' time investment or initial endowments. In the context of my model, the differences in total factor productivity can incorporate physiological or social differences in the trajectory of literacy accumulation that determine the efficiency of inputs. It can also reflect the role of other inputs that are not included in this model, like the quality of schooling and peers, which might play an increasingly important role as the child transitions to teenage years. An increasing amount of evidence suggests that boys and girls are differentially sensitive to the quality of neighbourhoods and schooling (Autor et al., 2020, Chetty et al., 2016).

Figure 4: Productivity estimates by age

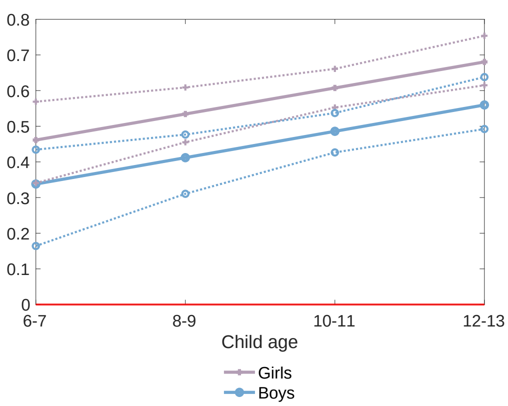
(a) Active time productivity, $\rho_t^{j,m}$



(b) Lag literacy productivity, $\rho_t^{j,s}$



(c) Total factor productivity, $\ln R_t^j$



Notes: This figure displays the estimates of the production function of log literacy for boys and girls. All panels additionally display 90% confidence intervals computed with 500 bootstrap iterations of the estimation procedure. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

4.2 Wage and non-labour income processes

The decisions of mothers regarding the time investment in their children depend on the parameters of their budget constraints.[13] In the model, higher available non-labour income would have a positive income effect on the time mothers spend with children, while wages represent an opportunity cost of time invested in children. Therefore, the mother's wage and non-labour income processes are important for modelling the mother's time investment decisions.

Even though the mother's hourly wages are measured in the data (see Section 3), the selection of mothers into employment represents a concern for estimating the wage process while using only observed positive hourly wages. To address the bias from endogenous selection into work, when estimating wage equations outside of the model, I use a Heckman selection correction with the number of siblings, available non-labour income and child's gender as excluded instruments.[14] I estimate the following wage process for mothers using the Heckman selection correction model

$$ln(w_{it}^{m*}) = \omega_0 + \omega_1 t + \omega_2 age_{it}^m + \omega_3 age_{it}^{m\ 2} + \omega_4 Educ_i^m + u_{it}^w, \tag{12}$$

where age_{it}^m is the mother's age, the child is of age t, and $Educ_i^m$ is equal to one if the mother has obtained a college degree before the child is 14-15. The wage shock is assumed to be IID normally distributed $u_{it}^w \sim N(0,\sigma^w)$. To estimate the wage shock distribution, I predict the wage equation's residuals and compute their variance.

The non-labour income process for the mother is modelled as a truncated latent variable process. I estimate separate non-labour income processes for single mothers and mothers in two-parent households. The latent non-labour income process is described by

$$\tilde{N}I_{it} = n_0 + n_1 t + n_2 age_{it}^m + n_3 age_{it}^{m\ 2} + n_4 Educ_i^m + n_5 nsib_{i,t} + u_{it}^n, \tag{13}$$

where the non-labour income shock follows an IID normal distribution $u_{it}^n \sim N(0,\sigma^n)$. The actual

[13]For example, Gennetian, Duncan, Fox, Magnuson, Halpern-Meekin, Noble, and Yoshikawa (2022) show that mothers increase learning activities with their children in response to an increase in non-labour income. Morrissey (2023) show that time investment in children is affected by changes to minimum wage laws.

[14]Appendix I shows the estimates for wage and non-labour income processes controlling for the gender of the child. The gender of the child is not a significant predictor of wage or non-labour income.

non-labour income process for all periods is given by

$$NI_{it} = max\{0, \tilde{NI}_{it}\}. \tag{14}$$

I estimate the non-labour income process using a Tobit regression model censored from below at zero values.

Column 1 of Table 2 shows that mothers' log hourly wages depend positively on their own age and education. Column 2 of Table 2 shows that the mother's education is positively related to the non-labour income levels in two-parent households, but negatively in single-parent households. Mothers who have more children have higher non-labour income values in both types of families. This heterogeneity in wage and non-labour income processes with respect to mother's characteristics, along with shocks to the wage and non-labour income, introduce variation in optimal labour supply decisions by mothers according to the Equation (11).

	Log mother wage	Non-lab. income - single	Non-lab. income - two parent
Child age	0.00	-37.70*	-24.38
	(0.01)	(8.03)	(16.19)
Mom age	0.05*	-8.85	65.81
	(0.02)	(6.60)	(42.23)
Mom age sqr.	-0.00*	0.11	-0.57
	(0.00)	(0.08)	(0.51)
Mom college	0.34*	-54.00*	408.48*
	(0.02)	(18.87)	(41.74)
Number of sib.		64.70*	74.66*
		(7.08)	(21.28)
Constant	1.72*	381.41*	-617.87
	(0.31)	(138.76)	(868.51)
SD shock	0.46	289.23	1160.62
	(0.01)	(30.24)	(52.45)
N	6474	1804	6759

Table 2: Estimates of Wage and Non-labour Income Processes

Notes: This table displays the estimates of the log hourly wage process and non-labour income process for mothers. Standard errors are obtained with 500 bootstrap iterations of the estimation procedure. The log hourly wage process is estimated using the Heckman selection correction, instrumenting for maternal employment with the number of children, non-labour income, and the child's gender. The non-labour income process is estimated using a Tobit regression model censored from below at zero values. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income. Significance level: * 5%.

4.3 Preference parameters and model fit

Conditional on the known parameters of exogenous processes, the preference parameters are identified from household choices. I estimate the mother's preferences for consumption $\alpha^{j,c}$ and human capital $\alpha^{j,s}$, time penalty for other children θ^{j}_{sib}, and the share of family income available to mothers b^{j}_{NI} using the Method of Simulated Moments. I assume the patience parameter to be $\beta = 0.95^2 = 0.9025$ to accommodate the biennial structure of the data. To allow for the flexibility of differences in preferences across boys and girls, I estimate them separately by gender. My estimation procedure involves several substeps. First, I draw wage and non-labour income shocks from the estimated IID normal distributions. Second, I draw initial skill and period skill shocks from the estimated gender-specific distribution. Third, I compute optimal investments and choices and summarize them using moments. Finally, I compare these with the data moments using a weight matrix, estimated as an inverse covariance matrix of data moments computed using bootstrap resampling from the original data.

All parameters of the model jointly determine household choices. However, some moments are more sensitive to specific parameters and are leveraged in the estimation. Consumption preferences are pinned down by matching the moments describing labour supply decisions, including the mean and standard deviation of the mothers' work hours and employment rate every period. Preferences for human capital vs leisure depend on the mothers' choices of active time with children vs leisure, summarized by the mean and standard deviation of the mothers' active time investment every period. Finally, I match unconditional moments including the correlations of active time investment and labour supply with the mothers' wages, non-labour income, and the number of siblings.[15]

Table 3 shows the estimated preference parameters for the households of boys and girls along with the bootstrap standard errors. Mothers of boys have a higher point estimate for consumption preferences and a lower point estimate for human capital preference, even though the differences are not statistically significant. While there is some evidence that families in de-

[15]See Appendix D for the results of a Monte Carlo exercise estimating preference parameters using the Method of Simulated Moments.

veloped and developing countries have a preference for sons[16], other work based on parental self-reporting highlights that parents do not have gender-specific preferences for human capital (Baker and Milligan, 2016), or they have higher educational aspirations for their daughters (Cobb-Clark and Moschion, 2017, Foley, 2019), which might reflect higher parental preferences for the human capital of their female children. My estimates are consistent with parents not having significant gender-related differences in preferences for human capital.

The estimated model fits the trends in mothers' active time investment and labour force participation observed in subsamples of boys and girls. Figure 5 shows that the model predicts the decline in the mother's time investment with the child's age, driven by the decrease in the productivity of the mother's time, and declining returns to investment. The model also explains the corresponding increase in mothers' labour supply both at the extensive (share of employed mothers) and intensive margins (average weekly labour hours).

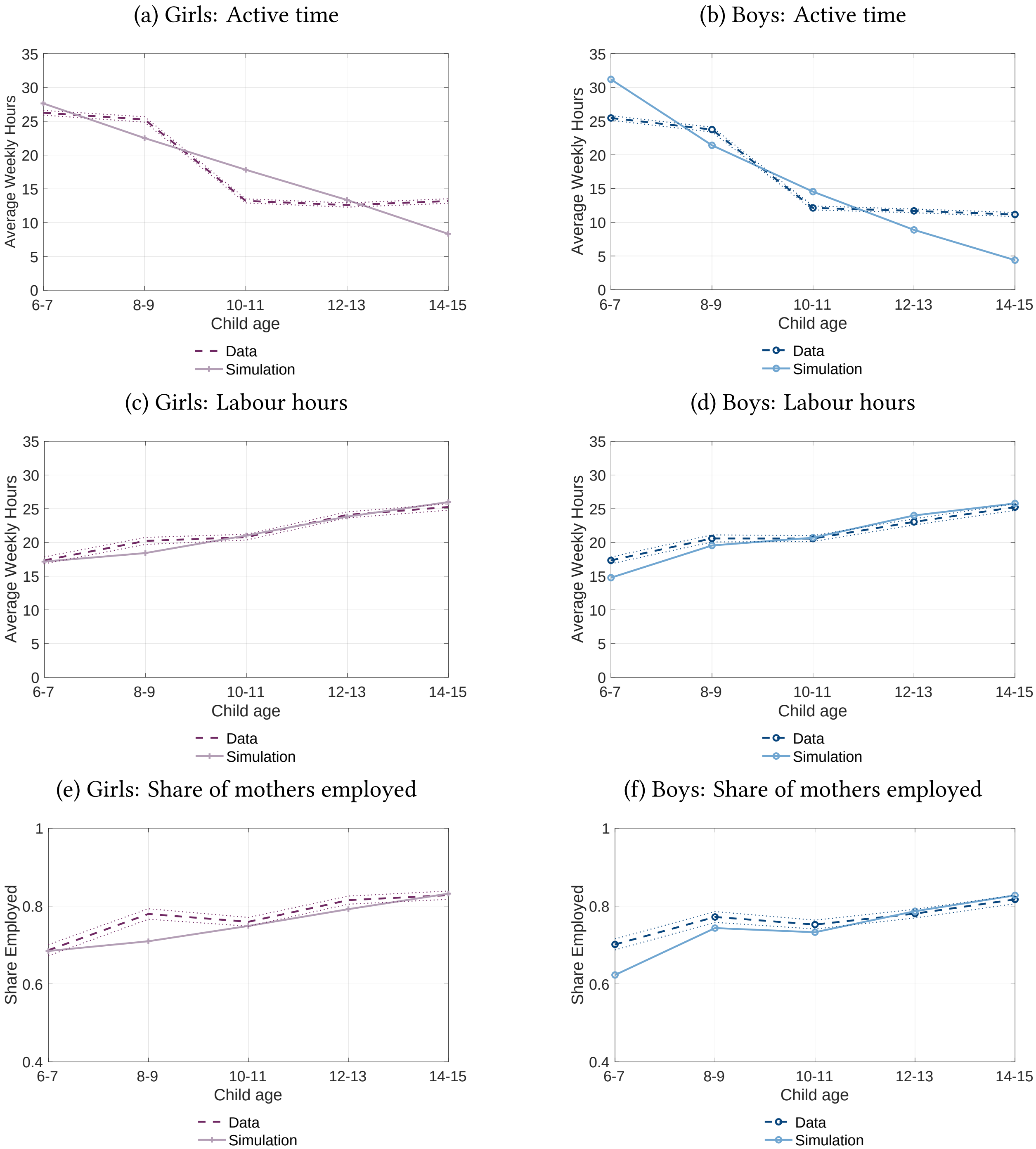
[16]For example, Dahl and Moretti (2008) show that having a first-born daughter in the U.S. is associated with a lower likelihood of becoming and staying married, as well as higher total number of children the family will choose to have. Barcellos, Carvalho, and Lleras-Muney (2014) show that first-born boys receive more parental investment compared to girls in India.

	Girls	Boys
Preference for consumption $\alpha^{j,c}$	0.033	0.079
	(0.063)	(0.068)
Preference for literacy $\alpha^{j,s}$	0.938	0.847
	(0.141)	(0.167)
Share of NI income b^j_{NI}	0.915	0.944
	(0.089)	(0.112)
Time penalty for a sibling Θ^j_{sib}	12.443	12.444
	(0.622)	(0.583)

Table 3: Estimates of Preference Parameters

Notes: This table displays the estimates of preferences for mothers of girls and boys. Standard errors are obtained with 500 bootstrap iterations of the estimation procedure. The preferences are estimated using the Method of Simulated Moments. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

Figure 5: Model fit - Targeted moments by child's age



Notes: This figure displays the model fit for average weekly time investment by mothers, labour supply, and share of employed mothers by child's age. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

5 Results

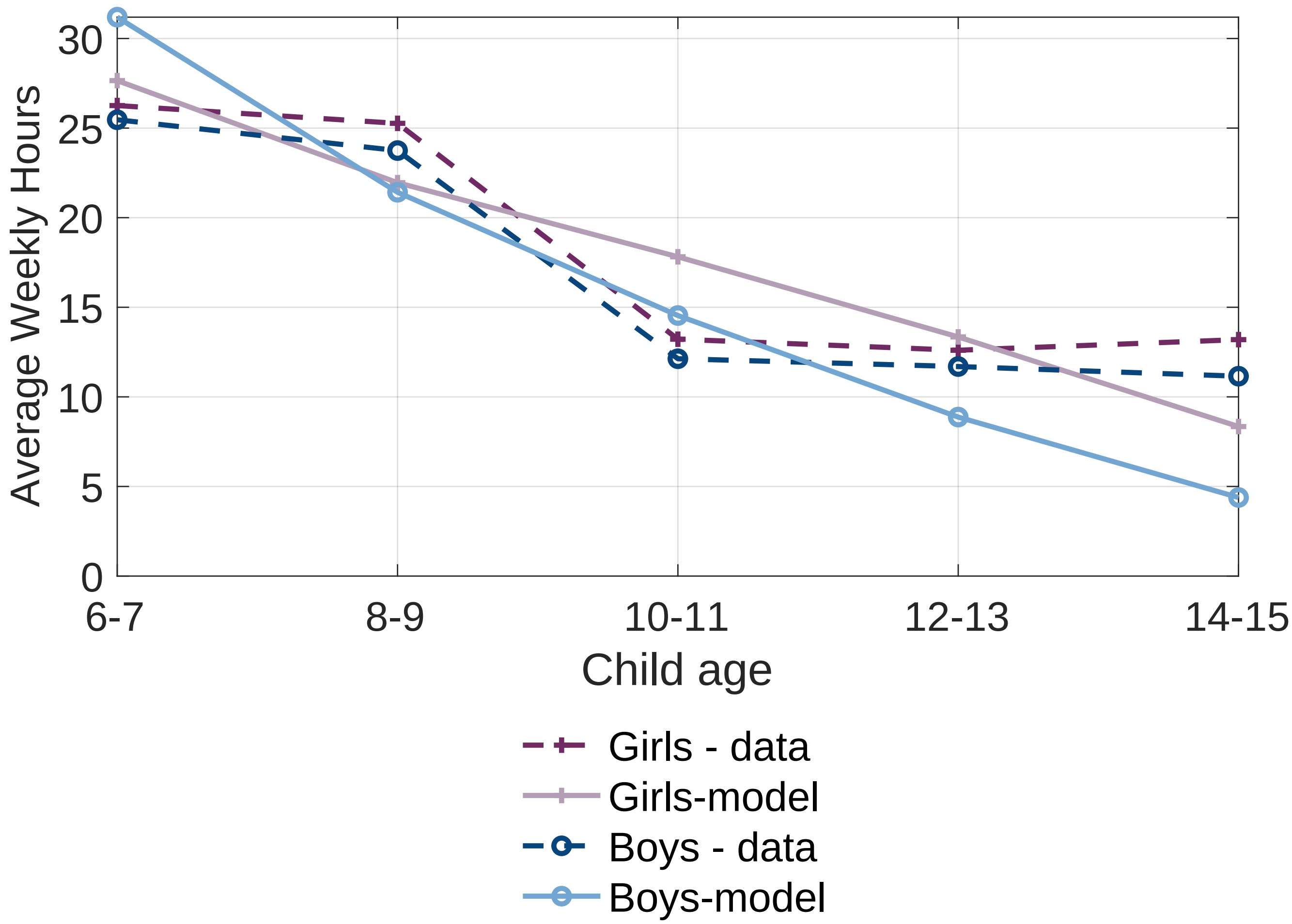
While gaps in parental investment decisions have been linked to deficits in human capital accumulation, the literature provides no definite answer on whether parental preferences for the human capital of boys and girls or for the time spent with children, parental constraints, or gender differences in the production function of human capital drive them. Distinguishing between behaviours due to various reasons can be a key input to policy responses that might alleviate the disadvantages introduced by adverse childhood circumstances (Almond, Currie, and Duque, 2018). This section discusses the implications of the estimated model for the gender gaps in parental investment and child literacy.

5.1 Predicting the active time investment gap

Panels (a) and (b) of Figure 5 show that the model of household choice predicts the average magnitudes and the decreasing patterns of parental time investment. Figure 6 additionally compares the average active time investment for sons and daughters in the data and in the baseline model. Appendix A presents the estimates reported in this subsection with the addition of 90% bootstrap confidence intervals. Point estimates of the model predict a positive female-favourable gap in parental investment development for children older than 10-11, when the investment gap in the data is statistically significant (see Figure 2b). At the same time, this predicted positive gender gap in parental investment is not statistically significant (see Panel (a) of Figure A.1).

To understand what drives the gender gap in parental time, I analyze the contribution of three factors to predicted gender gaps in parental investment. First, I quantify the role of the difference in the self-productivity of literacy captured by $\rho_t^{j,s}$ in Equation (2). Second, I focus on the role of the productivity of active time with mothers $\rho_t^{j,m}$ in Equation (2). Finally, I quantify the role of differences in preferences $\alpha^{j,s}$ and $\alpha^{j,c}$, access to family budget b_{NI}^j, and time penalties for siblings θ_{sib}^j.

Figure 6: Average active time investment by age for boys and girls, data and model



Notes: This figure displays the average weekly time investment by mothers in the data and in the baseline model for boys and girls. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

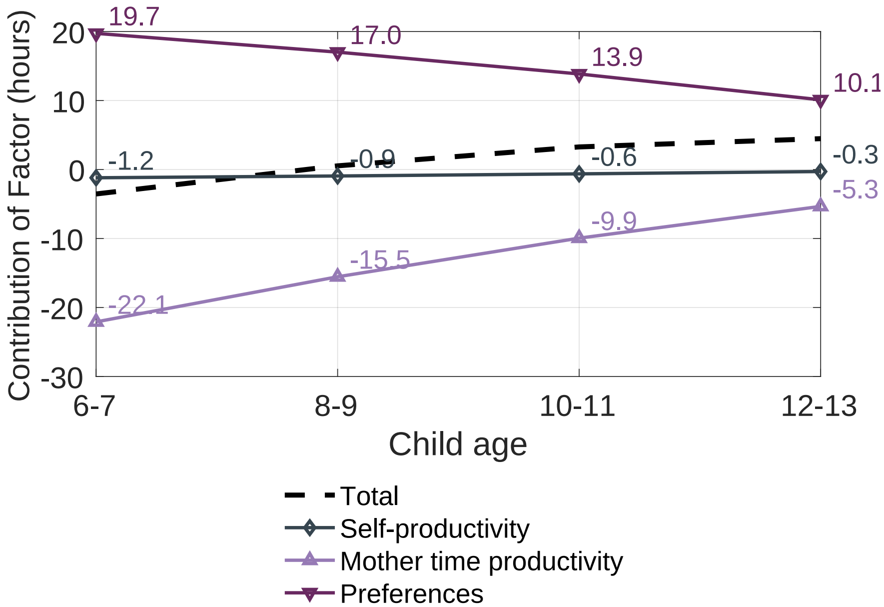
To quantify the contribution of various factors to the female-favourable gap in mother's time investment, I perform the Shapley-Owen-Shorrocks decomposition. The Shapley-Owen-Shorrocks decomposition ensures that the order in which the influence of the factors is shut down in the model does not matter for the decomposition results. Additionally, it allows interpretation of the contribution of any particular factor as a proportion of the increase in the gender gap in literacy attributed to this factor. I analyze the role of assigning production function parameters and preferences specific to girls to the baseline model for boys. To quantify the contribution of each factor, I perform the decomposition in all possible permutations of the order of decomposition and compute the average marginal contribution of each factor across all permutations.

Figure 7 shows the predicted effect of these various factors on the female-favourable gap in time investment.[17] Since boys have higher returns to the active time investment by mothers according to the point estimates of productivity (see Figure 4a), boys' advantage would encourage mothers to invest more time in sons compared to daughters, particularly for young children. As boys' time productivity advantage shrinks with age, the incentive to invest more in sons compared to daughters also decreases. On another hand, higher point estimates of mother's preference for the literacy of daughters would encourage mothers to invest more in daughters compared to sons, particularly at a young age when the marginal benefit of time investment is compounded due to the positive self-productivity of skills. The model's point estimates predict that higher preferences for literacy would dominate the role of productivity advantage as the child grows older, explaining the appearing female-favourable gender gap in time investment in teenagers older than 10-11.

Since the role of these various factors is not statistically significant (see Panels (b) to (d) of Figure A.1), my estimates would be consistent with some of the evidence on self-reported parental preferences which did not reveal significant differences between boys and girls. For example, Baker and Milligan (2016) find that parents' self-reported valuation of certain qualities of children like doing well at school do not differ significantly for sons and daughters. They also did not see significant gaps in whether children of different genders were wanted. It is possible that while parents do not have gender-specific preferences for the human capital of their

[17] The share of the contribution of each factor can be computed by dividing the effect of each individual factor by the total value.

Figure 7: Decomposition of the female-favourable gap in active time investment, hours



Notes: This figure displays the Shapley-Owen-Shorrocks decomposition of factors driving the gap in average weekly time investment by mothers, including preferences, productivity of mother's time, and self-productivity of log literacy. See Figure A.1 for the 90% confidence intervals obtained with 500 bootstrap iterations of the estimation procedure. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

children, they enjoy spending learning-oriented time with their daughters more, for example, if boys do not like being read to and misbehave. In Appendix H I show that there are no significant gender differences in the duration of time for which the child enjoys being read to according to mothers. However, there are some significant female-favourable gender gaps in the educational aspirations of mothers, which would be consistent with mothers highly valuing the literacy of their daughters.

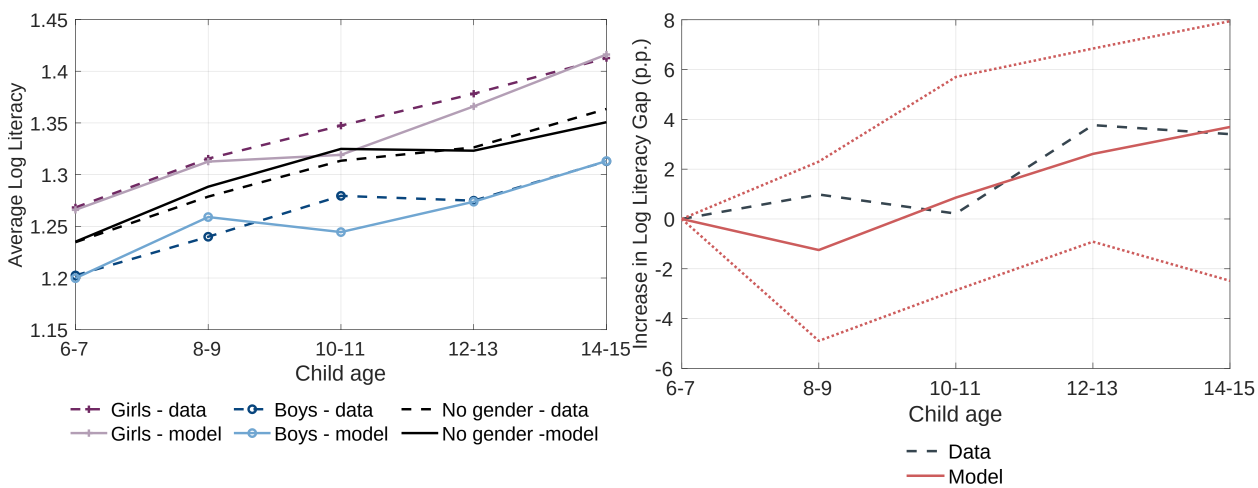
5.2 Predicting the expanding literacy gap

The main concern with differences in productive inputs in boys and girls is the potential detrimental effect of investment deficits on literacy accumulation in boys. Panel (a) of Figure 8 displays the evolution of the average literacy for boys and girls in the data and in the baseline model. It also shows the predicted levels of literacy for children if the model was estimated without allowing for gender-specific heterogeneity in endowments, preferences, and production function. This model predicts the same trajectory of skills for boys and girls and abstracts from differences in developmental trajectories. Panel (b) of Figure 8 illustrates the increase in the female-favourable gender gap in literacy observed in the data along with that predicted in the model with 90% confidence intervals. The model predicts that the female-favourable gender gap in literacy expands by 3.7 percentage points from 6.6 percent in children ages 6-7 to 10.3 percent in children ages 14-15. This is consistent with the increase from 6.6 to 10.0 percent observed in the data.

Figure 8: Model fit: female-favourable gap in log literacy

(a) Average log literacy by age

(b) Increase in female-favourable gap, p.p.



Notes: This figure displays the model fit for the average log literacy levels. Panel a) plots the average log literacy levels for boys, girls, and all children by child's age in the data and in the baseline model. Panel b) plots the increase in the female-favourable gap in log literacy levels by age in the data and in the baseline model. The figure includes the 90% confidence intervals for the model predictions obtained with 500 bootstrap iterations of the estimation procedure. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

To quantify the role of various drivers in the expansion of the gender gap in literacy, I perform the Shapley-Owen-Shorrocks decomposition for five factors. First, I analyze the contribution of gender differences in the TFP part of the production function R_t^j in Equation (2).

Second, I quantify the role of the difference in the self-productivity of literacy captured by $\rho_t^{j,s}$ in Equation (2). Third, I focus on the role of the mother's time $\rho_t^{j,m}$ in Equation (2). Fourth, I quantify the role of differences in preferences $\alpha^{j,s}$ and $\alpha^{j,c}$, access to family budget b_{NI}^j, and time penalties for siblings θ_{sib}^j. Finally, I account for the contribution of differences in the distribution of initial endowments of literacy μ_1^j and σ_1^j.

Figure 9 shows the predicted effect of these various factors on the change in the female-favourable gap in literacy compared to the initial (age 6-7) levels.[18] Appendix B additionally reports the 90% bootstrap confidence intervals for the estimates in this Subsection. Based on the point estimates of production function, assigning the female point values of the productivity of mother's time would reverse the gender gap in literacy over time and decrease it by over 5 percentage points. Assigning female values of self-productivity to boys would also reverse the gender gap by 11.1 percentage points. This would be driven by both lower productivity of inputs and by mothers decreasing their active time investment to reflect the decreasing investment returns. Because inputs are slightly more productive for boys than they are for girls, input productivity differences alone cannot explain the expansion of the female-favourable gap in literacy.

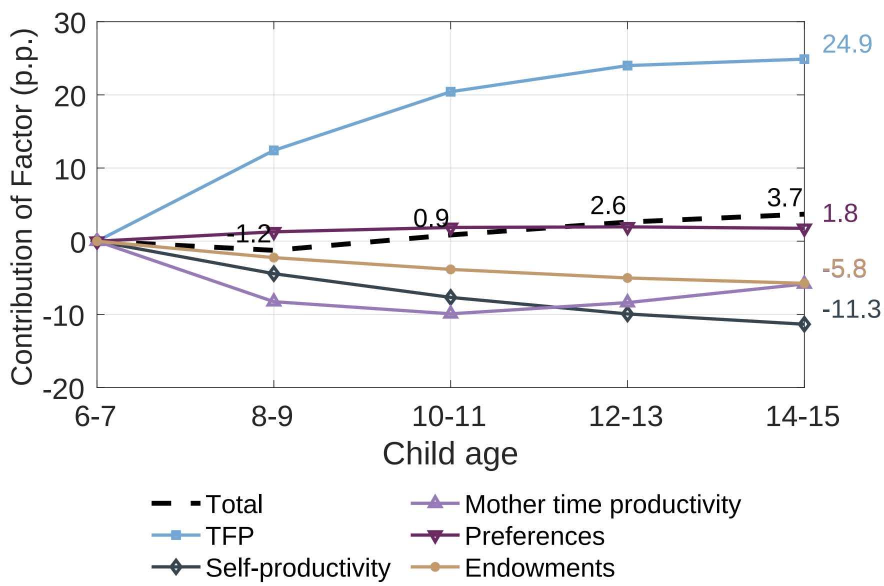
The contribution of preferences, which slightly favour girls based on point estimates of the model, is positive but small, predicting a 1.7 percentage point expansion of the gender gap. Together with the minor estimated role of differences in active time productivity, these results suggest that the drivers behind differences in mothers' aggregate active time investment play a limited role in explaining the expansion of the female-favourable gap in literacy.

The main factor contributing to the expansion of the female-favourable gap in literacy is captured by the total factor productivity of the production function. The female advantage in TFP would drive the gender gap to increase by over 20 percentage points by the age of 14-15 if other factors did not counterbalance this effect. With both the productivity of time investment and the magnitude of time investment declining with the child's age, a growing share of the growth in children's literacy scores is not explained by mothers' time investment or initial endowments. This share is also higher for girls than it is for boys. This role of TFP can result

[18]The share of the contribution of each factor can be computed by dividing the effect of each individual factor by the total value.

from multiple factors, including genetic predisposition to different trajectories of brain maturation (Lim, Han, Uhlhaas, and Kaiser, 2015), the importance of factors like neighbourhoods and schools (Autor, Figlio, Karbownik, Roth, and Wasserman, 2020, Chetty, Hendren, Lin, Majerovitz, and Scuderi, 2016), differences in behavioral skills (Bertrand and Pan, 2013), and differences in controlling and encouraging parenting practices across boys and girls (Rogers, Theule, Ryan, Adams, and Keating, 2009). Overall, my results suggest that the main drivers of the expansion of female-favourable gaps in literacy are productivity differences unrelated to parental aggregate time investment.

Figure 9: Decomposition of the expansion in the female-favourable gap in literacy



Notes: This figure displays the Shapley-Owen-Shorrocks decomposition of factors driving the increase in the female-favourable gap in log literacy score, including preferences, productivity of mother's time, self-productivity of log literacy, TFP, and initial endowments. See Figure B.1 for the 90% confidence intervals obtained with 500 bootstrap iterations of the estimation procedure. The sample includes children in LSAC K-cohort with non-missing measures of the mother's active time investment and labour supply, positive measures of active time investment, who appeared in the survey at least twice, and non-negative non-labour income.

6 Conclusion

This paper leverages a structural model of a mother's time investment choice to quantify the contribution of gender differences in the production function of skill, initial endowments, and preferences to the expansion of female-favourable literacy deficits in school-aged children in Australia. I find that the gaps in maternal investment are mainly explained by the mother's preference for the literacy of daughters or the time spent with daughters, while the productivity of time investment benefits boys, however, the estimates are not statistically significant. Importantly, my estimates suggest that gender differences in mother's time investment do not play a big role in explaining the expansion of skill deficits in children. One explanation behind this limited role is that parental time was shown to be particularly important for young children, while my analysis is focused on children ages 6 to 15. My estimates suggest that for older children exploring the role of other factors like schooling, peer groups, and neighbourhoods might be a promising direction for future research.

References

Francesco Agostinelli and Matthew Wiswall. Estimating the technology of children's skill formation. Technical report, National Bureau of Economic Research, 2016.

Douglas Almond, Janet Currie, and Valentina Duque. Childhood circumstances and adult outcomes: Act ii. Journal of Economic Literature, 56(4):1360–1446, 2018.

David Autor, David N Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. Males at the tails: How socioeconomic status shapes the gender gap. Technical report, National Bureau of Economic Research, 2020.

Michael Baker and Kevin Milligan. Boy-girl differences in parental time investments: Evidence from three countries. Journal of Human Capital, 10(4):399–441, 2016.

Silvia Helena Barcellos, Leandro S Carvalho, and Adriana Lleras-Muney. Child gender and parental investments in india: Are boys and girls treated differently? American Economic Journal: Applied Economics, 6(1):157–189, 2014.

Marianne Bertrand and Jessica Pan. The trouble with boys: Social influences and the gender gap in disruptive behavior. American economic journal: applied economics, 5(1):32–64, 2013.

Costanza Biavaschi, Corrado Giulietti, and Klaus F Zimmermann. Sibling influence on the human capital of the left-behind. Journal of Human Capital, 9(4):403–438, 2015.

Francesca Borgonovi, Álvaro Choi, and Marco Paccagnella. The evolution of gender gaps in numeracy and literacy between childhood and adulthood. 2018.

Anne Ardila Brenøe and Shelly Lundberg. Gender gaps in the effects of childhood family environment: Do they persist into adulthood? European Economic Review, 109:42–62, 2018.

Ylenia Brilli. Mother's time allocation, childcare, and child cognitive development. Journal of Human Capital, 16(2):233–272, 2022.

Claudia Buchmann and Thomas A DiPrete. The growing female advantage in college completion: The role of family background and academic achievement. American sociological review, 71 (4):515–541, 2006.

Elizabeth M Caucutt and Lance Lochner. Early and late human capital investments, borrowing constraints, and the family. Journal of Political Economy, 128(3):1065–1147, 2020.

Elizabeth M Caucutt, Lance Lochner, Joseph Mullins, and Youngmin Park. Child skill production: Accounting for parental and market-based time and goods investments. Technical report, National Bureau of Economic Research, 2020.

Raj Chetty, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. Childhood environment and gender gaps in adulthood. American Economic Review, 106(5):282–88, 2016.

Deborah A Cobb-Clark and Julie Moschion. Gender gaps in early educational achievement. Journal of Population Economics, 30:1093–1134, 2017.

Christopher Cornwell, David B Mustard, and Jessica Van Parys. Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. Journal of Human resources, 48(1):236–264, 2013.

Flávio Cunha. Gaps in early investments in children. Manuscript, Dept. Econ., Univ. Pennsylvania, 2014.

Flavio Cunha and James J Heckman. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. Journal of human resources, 43(4):738–782, 2008.

Gordon B Dahl and Enrico Moretti. The demand for sons. The review of economic studies, 75(4): 1085–1120, 2008.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. Review of Economic Studies, 81(1):137–185, 2014.

Doris R Entwisle, Karl L Alexander, and Linda S Olson. Early schooling: The handicap of being poor and male. Sociology of education, 80(2):114–138, 2007.

Xiaodong Fan, Hanming Fang, and Simen Markussen. Mothers' employment and children's educational gender gap. Technical report, National Bureau of Economic Research, 2015.

Mario Fiorini and Michael P Keane. How the allocation of children's time affects cognitive and noncognitive development. Journal of Labor Economics, 32(4):787–836, 2014.

Kelly Foley. The gender gap in university enrolment: Do parents play a role beyond investing in skills? Canadian Journal of Economics/Revue canadienne d'économique, 52(2):441–489, 2019.

Lisa A Gennetian, Greg Duncan, Nathan A Fox, Katherine Magnuson, Sarah Halpern-Meekin, Kimberly G Noble, and Hirokazu Yoshikawa. Unconditional cash and family investments in infants: Evidence from a large-scale cash transfer experiment in the us. Technical report, National Bureau of Economic Research, 2022.

Elisabeth Gugl and Linda Welling. Time with sons and daughters. Review of Economics of the Household, 10:277–298, 2012.

Luigi Guiso, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. Culture, gender, and math. Science, 320(5880):1164–1165, 2008.

Sunhee Kim and Gregory Camilli. An item response theory approach to longitudinal analysis with application to summer setback in preschool language/literacy. Large-scale assessments in Education, 2(1):1–17, 2014.

Victor Lavy and Edith Sand. On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases. Technical report, National Bureau of Economic Research, 2015.

Sol Lim, Cheol E Han, Peter J Uhlhaas, and Marcus Kaiser. Preferential detachment during human brain development: age-and sex-specific structural connectivity in diffusion tensor imaging (dti) data. Cerebral Cortex, 25(6):1477–1489, 2015.

Shelly J Lundberg. Father absence and the educational gender gap. 2017.

Stephen Machin and Tuomas Pekkarinen. Global sex differences in test score variability. Science, 322(5906):1331–1332, 2008.

Gurian Michael and Kathy Stevens. The minds of boys: Saving our sons from falling behind in school and life. Wiley+ ORM, 2010.

Taryn W Morrissey. The minimum wage and parent time use. Review of Economics of the Household, 21(3):1043–1062, 2023.

Joseph Mullins. Designing cash transfers in the presence of children's human capital formation. Job Market Paper.[235], 2019.

Cheti Nicoletti and Valentina Tonei. Do parental time investments react to changes in child's skills and health? European Economic Review, 127:103491, 2020.

Cheti Nicoletti, Almudena Sevilla, and Valentina Tonei. Gender stereotypes in the family. 2022.

Matthias Pollmann-Schult. Sons, daughters, and the parental division of paid work and housework. Journal of Family Issues, 38(1):100–123, 2017.

Joseph Paul Robinson and Sarah Theule Lubienski. The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. American Educational Research Journal, 48(2):268–302, 2011.

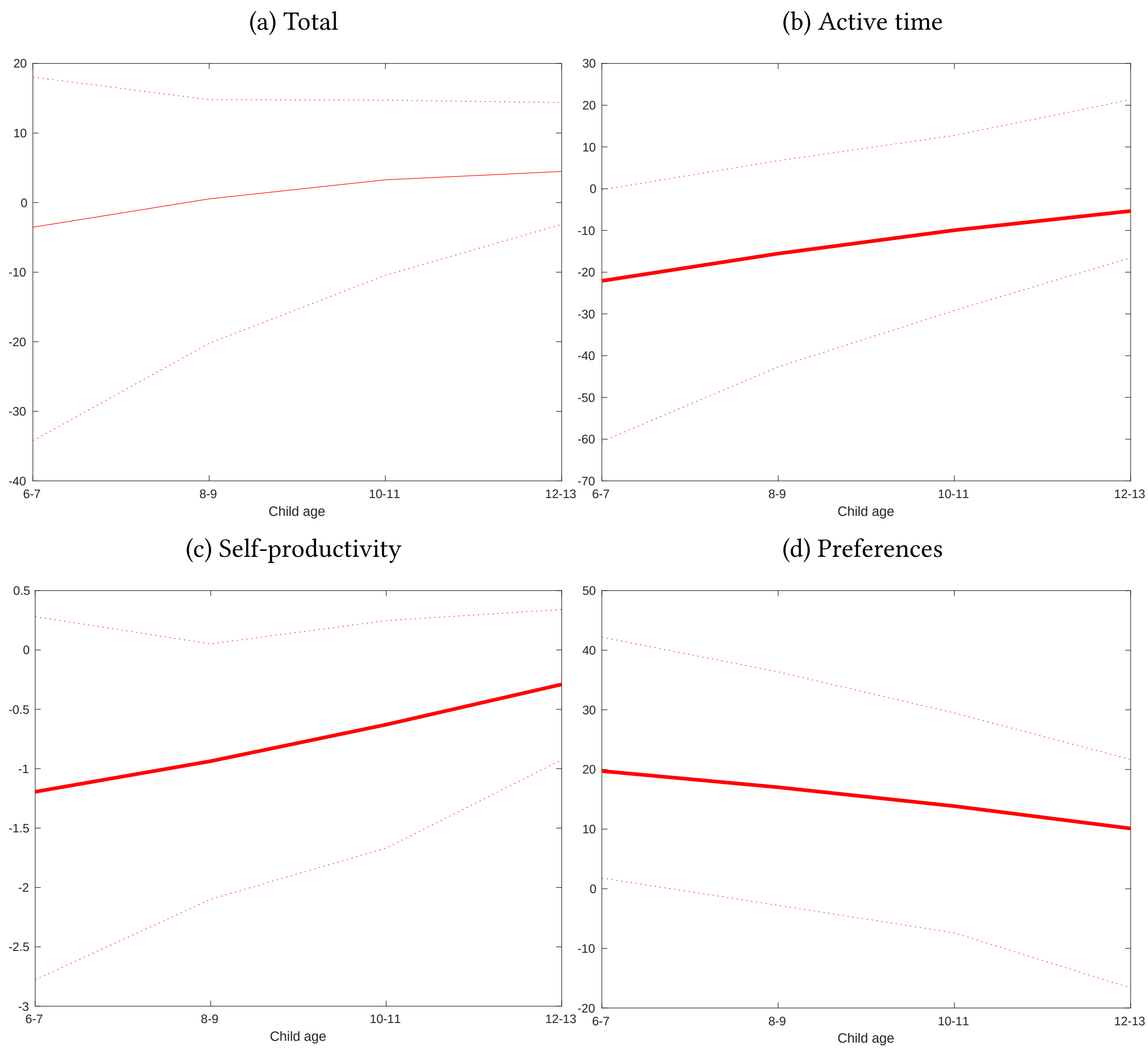
Núria Rodríguez-Planas and Natalia Nollenberger. Let the girls learn! it is not only about math... it's about gender social norms. Economics of Education Review, 62:230–253, 2018.

Maria A Rogers, Jennifer Theule, Bruce A Ryan, Gerald R Adams, and Leo Keating. Parental involvement and children's school achievement: Evidence for mediating processes. Canadian Journal of School Psychology, 24(1):34–57, 2009.

Florian Schulz. Mothers', fathers' and siblings' housework time within family households. Journal of Marriage and Family, 83(3):803–819, 2021.

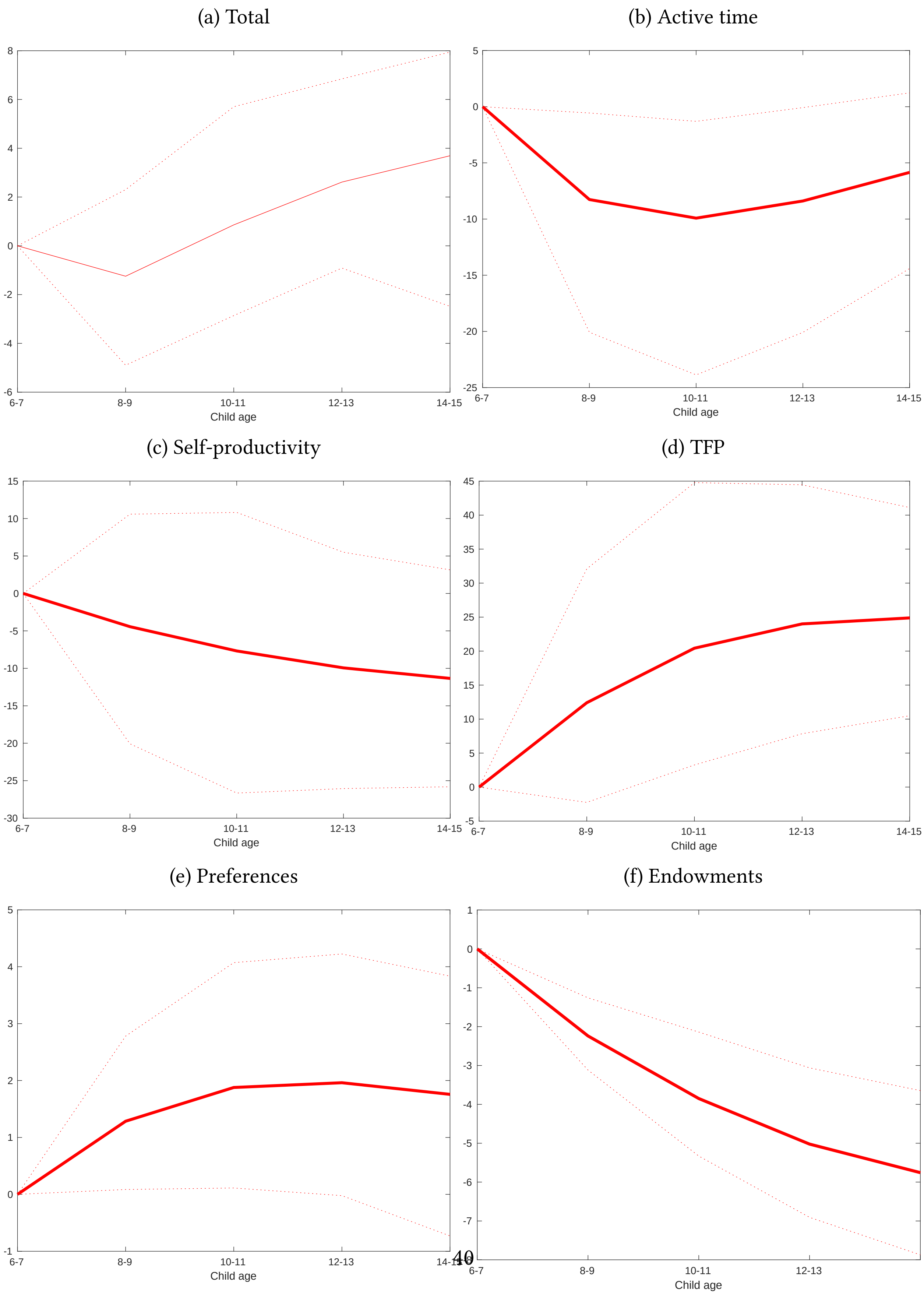
A Decomposition of gaps in active time investment - bootstrap standard errors

Figure A.1: Decomposition of gaps in time investment



B Decomposition of gaps in log literacy - bootstrap standard errors

Figure B.1: Decomposition of expansion of the gaps in log literacy

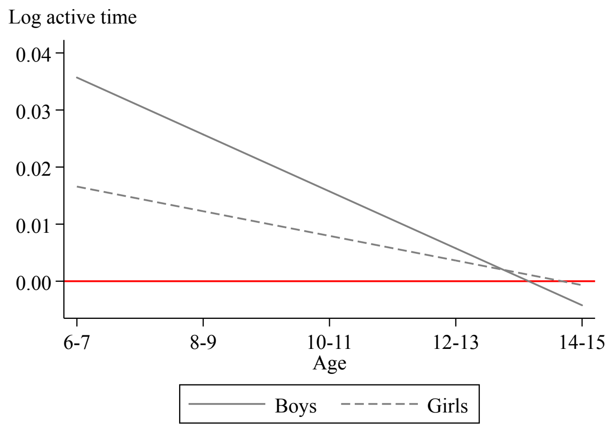


C Production function and measurement error in literacy score

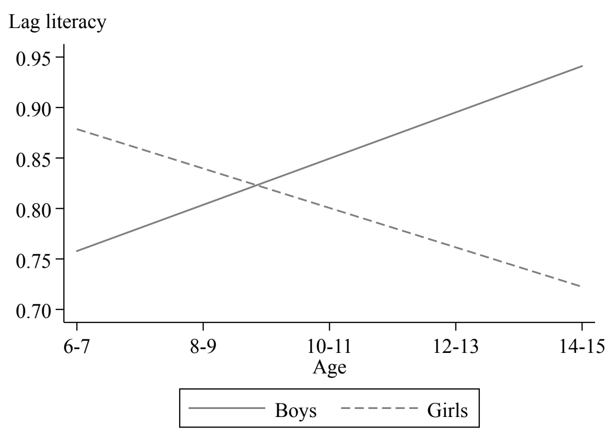
A measurement error in the literacy score can bias the estimates of my production function. To address this concern, I estimate a production function linear in parameters while instrumenting for the lag of teacher-reported log literacy score with an alternative measure of children's cognitive achievement (Agostinelli and Wiswall, 2016). The LSAC does not consistently collect alternative measures of absolute progress in literacy. However, it asks mothers about their child's achievement in reading and school compared to other children in their class. I use the mother's report on the child's reading and overall school achievement as an alternative measure of literacy. Figure C.1 presents the estimates of productivity by age. After adjusting for the measurement error in log literacy active time investment is still slightly more productive for boys than it is for girls, and the magnitude of estimated productivity is comparable to the baseline estimates. The female TFP remains higher than the male for children older than 8-9.

Figure C.1: Productivity estimates adjusted for measurement error in literacy score, by age

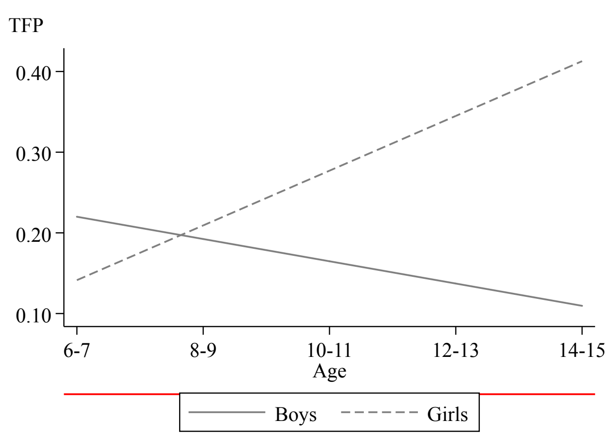
(a) Active time productivity, $\rho_t^{j,m}$



(b) Lag literacy productivity, $\rho_t^{j,s}$



(c) Total factor productivity, $\ln R_t^j$

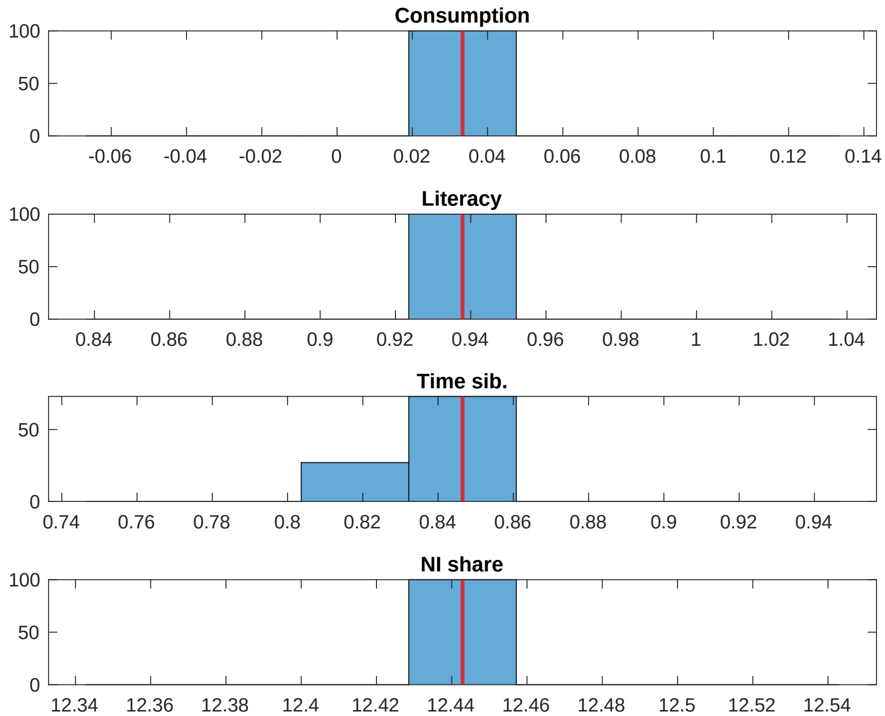


D Monte Carlo estimation of the preference parameters

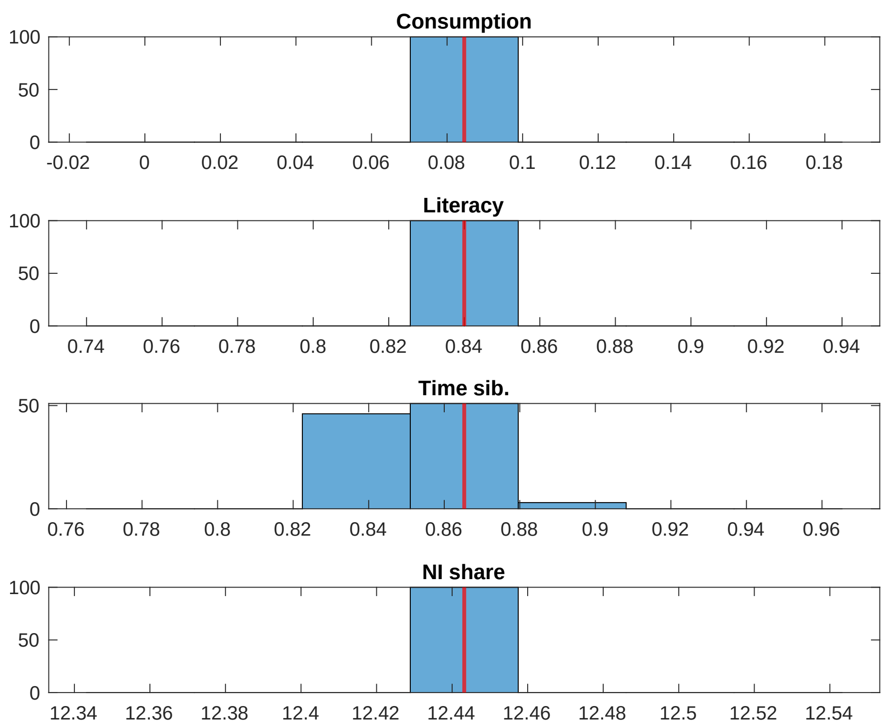
This section shows the results of a Monte-Carlo exercise estimating preference parameters conditional on the estimates of production function, initial skill distribution, and income processes. The estimates are obtained by simulating 100 datasets based on preference parameters reported for boys and girls in Subsection 4.3. Figure D.1 shows that parameters of preferences are precisely estimated for subsamples of boys and girls.

Figure D.1: Monte-Carlo preference estimation - 100 simulations.

(a) Girls



(b) Boys



E Time investment categories

Time spent in the following activities is counted as active time investment by mothers if the mother is present during the activity:

Age 6-7:

Eating and drinking;

Bathe, dress, hair care, health care;

Do nothing, bored/restless;

Held, cuddled, comforted, soothed;

Read a story, talk/sing, talked/sung to;

Reading, looking at books by self;

Quiet free play;

Active free play;

Helping with chores/jobs;

Visiting people, special event, party;

Organized sport/physical activity;

Other organized lessons/activities;

Walk for travel or for fun;

Ride bicycle, trike etc (travel or fun);

Taken places with adult (e.g. shopping);

Age 8-9:

Eating, drinking, being fed;

Bathe, dress, hair care, health care;

Being hugged, comforted, etc.;

Read to or told a story;

Reading or looking at a book by self;

Quiet free play;

Active free play;

Helping with chores;

Visiting people, special events, outing;

Organized sport/physical activity;

Other organized lessons/activities;

Walking;

Riding bike etc.;

Taken places with adults;

Age 10-11:

Eating/drinking;

Personal/Health care;

Dentist, Doctor, Chiropractor, Physio etc.;

Cooking, meal preparation, making lunch, setting table for others;

Washing dishes, stacking and emptying dishwasher;

Taking care of siblings, and other children;

Active Activities;

Organized team sports and training i.e football, basketball, netball etc.;

Organized individual sports i.e. swimming, dancing, martial arts, etc.;

Ball games, riding a bike, scooter, skateboard, skipping, running, games and other free activities;

Taking Pet for a walk;

Scouts, girl guides, etc.;

Shopping;

Going out to museums, cultural events, fairs, community events, church etc.;

Cinema;

Live Sporting Events;

Non Active Activities;

Private music, language, religion, tutoring lessons;

Reading or being read to for leisure;

Board or card games, puzzles, toys, arts and crafts, etc.;

Non-Active Club Activities i.e. Chess Club;

Homework (not on computer) including music practice;

Computer for homework - internet;

Computer for homework - not internet;

Communication;

Talking face to face;

Travel by foot Travel by bike, scooter, skateboard;

Age 12-13 and 14-15 :

Eating/drinking;

Doctor;

Dentist;

Physiotherapist/Chiropractor;

Medical/Health care nec.;

Food/drink preparation;

Food/drink clean up;

Taking care of siblings;

Organized team sports and training;

Organized individual sport and training;

Unstructured active play;

Walking pets/playing with pets;

Active club activities;

Shopping;

Going out to a concert, play, museum, art gallery, community or school event , an amusement park etc.;

Religious activities / ritual ceremonies;

Attending live sporting events;

Active activities nec.;

Private music lessons/practice, academic tutoring;

Playing musical instruments or singing for leisure;

Reading or being read to for leisure;

Unstructured non-active play;

Non-active club activities;

Doing homework (not via electronic devices);

Non-active activities nec.;

Doing homework;

Creating/maintaining websites (excluding social networking profile);

General application use (e.g. Microsoft Office; excluding homework);

Talking face-to-face (in person not via electronic devices);

Non-verbal interaction (e.g. cuddles);

Travel by foot;

Travel by bike, scooter, skateboard etc.;

F Number of siblings and input productivity

The limitation of LSAC data is that information about parental investment is collected for one child per household, while the majority of children in the sample have siblings. In the presence of different caregiving roles, boys and girls might be affected differently by the presence of siblings (Biavaschi, Giulietti, and Zimmermann, 2015, Schulz, 2021). For example, it is possible that since girls share more of the household responsibilities than boys, the productivity of inputs in their human capital accumulation will be diluted with the presence of siblings to take care off.

Table F.1 shows the estimates of the production function for boys and girls $j = \{b,g\}$ obtained with the non-linear least squares allowing for the number of siblings to affect the productivity of inputs.

$$\ln S_{t+1} = exp(\gamma^j + \gamma^j_{age} t + \gamma^j_{nsib} nsib_{it} + \nu^j_{it}) + exp(\gamma^{j,m} + \gamma^{j,m}_{age} t + \gamma^{j,m}_{nsib} nsib_{it}) \ln \tau^m_t + \\ exp(\gamma^{j,s} + \gamma^{j,s}_{age} t + \gamma^{j,s}_{nsib} nsib_{it}) \ln S_t \quad (15)$$

Column (1) presents the estimated γ for boys under the assumption that siblings do not affect the productivity of inputs similarly to the baseline model. Column (2) presents the estimated γ in Equation (15) along with the p-value for the Wald test on joint significance of coefficients γ^j_{nsib},

$\gamma^{j,m}_{nsib}$, and $\gamma^{j,s}_{nsib}$. It can be seen that controlling for the number of siblings does not significantly affect the estimated productivity of inputs. The hypothesis that coefficients related to siblings are equal to zero cannot be rejected at 10% significance level. Columns (3) and (4) report similar estimates for the subsample of girls where the hypothesis that coefficients related to siblings are equal to zero also cannot be rejected at 10% significance level.

Table F.1: Production function estimates controlling for the number of siblings

	Boys		Girls	
	(1)	(2)	(3)	(4)
Cons.	0.19	0.15	0.31*	0.26*
	(0.12)	(0.12)	(0.12)	(0.12)
Age	0.07*	0.07*	0.07*	0.08*
	(0.03)	(0.03)	(0.03)	(0.03)
Lit. - cons.	-0.25	-0.18	-0.29	-0.23
	(0.14)	(0.15)	(0.15)	(0.15)
Lit. - age	-0.06	-0.06	-0.07*	-0.08*
	(0.03)	(0.03)	(0.04)	(0.04)
Act time - cons	-2.75*	-2.83*	-4.26*	-3.71*
	(0.70)	(0.83)	(2.11)	(1.58)
Act. time - age	-0.38	-0.39	-0.17	-0.21
	(0.20)	(0.20)	(0.52)	(0.40)
Num. sib.		0.03		0.03
		(0.02)		(0.02)
Lit. - num. sib.		-0.05		-0.03
		(0.03)		(0.03)
Act. time - num. sib.		0.05		-0.20
		(0.21)		(0.40)
Child-year obs.	1948	1948	2055	2055
Wald test (num. sib.) p-value		0.37		0.56

G Full sample summary

Table 1 summarizes the household characteristics of children between ages 6 and 15 in the K-cohort before sample restrictions are introduced. The main source of attrition for children ages 6-7 and 8-9 in the data are the time use diaries and teacher-reported literacy measures. Only 68% of mothers of children ages 8-9 have participated in time diaries collection. Additionally, only around 80% of teachers submit their questionnaires, which limits the number of children for whom the teacher-reported literacy score is available.

Table G.1: Summary of household characteristics by child's age

	Ages 6-7			Ages 8-9			Ages 10-11			Ages 12-13			Ages 14-15		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Mothers hourly wage 2003 dollars				2621	23.00	13.12	2588	23.34	12.46	2516	23.88	12.78	2304	23.98	12.31
Non-labour income including spousal wage				3723	1089.46	999.24	3560	1068.54	1146.08	3282	1080.35	1156.22	2909	1041.38	1022.20
Weekly work hours	4420	16.46	16.45	4284	19.73	16.57	4107	21.22	16.67	3877	23.62	16.87	3398	25.36	17.10
Mother employed	4424	0.66	0.47	4287	0.74	0.44	4108	0.76	0.42	3944	0.80	0.40	3533	0.83	0.38
Mother college+	4462	0.36	0.48	4330	0.37	0.48	4164	0.37	0.48	3956	0.38	0.48	3538	0.39	0.49
Mother age	4423	36.90	5.26	4286	38.96	5.27	4111	41.07	5.32	3883	43.20	5.17	3456	45.35	5.15
Num. of siblings	4463	1.58	1.04	4331	1.64	1.07	4164	1.63	1.07	3951	1.59	1.08	3527	1.50	1.04
Two-parent household	4463	0.81	0.39	4331	0.79	0.41	4164	0.76	0.42	3956	0.74	0.44	3538	0.73	0.45
Share of positive active time reports	1500	0.94	0.23	833	0.96	0.19	3989	0.89	0.31	3646	0.89	0.32	3074	0.81	0.39

H Child's gender and parental self-reported preferences

My results are suggestive of mothers spending more time with daughters because of slight female-favourable preferences for literacy or time spent with daughters. It is possible that mothers spend less time with their sons because they respond to boys enjoying the time spent in joint educational activities like reading less. When children are ages 4 to 5 and 6 to 7, mothers' reports how many minutes the child enjoys being read to. Table H.2 shows the estimates form the OLS regression of the number of minutes the child enjoys being read to on the child gender dummy and a set of controls. When the child is 4-5, no measures of teacher-reported literacy are available. Instead, column (2) controls for the child's Peabody Picture Vocabulary Test (PPVT) score, which measures children's receptive vocabulary, and mother-reported non-cognitive problems score.[19]

[19] The score is an age-standardized index based on the Strength and Difficulty assessment tool.

Columns (1) and (2) of Table H.2 show that mothers perceive young boys to have less preference for being read to compared to girls, even after controlling for children's skills. Columns (3) and (3) show that this gender difference is no longer significant for older children.

Table H.2: Gender gap in the number of minutes the child enjoys being read to

	Age 4-5		Age 6-7	
	(1)	(2)	(3)	(4)
Child female	2.82*	1.83*	0.33	0.14
	(0.54)	(0.56)	(0.28)	(0.33)
Family income	0.71*	0.27	0.24	0.11
	(0.33)	(0.35)	(0.15)	(0.17)
Mom coll+	4.16*	3.24*	2.44*	1.91*
	(0.62)	(0.65)	(0.32)	(0.37)
Num. sib.	-0.57*	-0.03	-0.72*	-0.77*
	(0.27)	(0.28)	(0.14)	(0.17)
Two parents	-0.55	-1.84	0.31	0.37
	(1.42)	(1.58)	(0.69)	(0.80)
PPVT score		2.57*		
		(0.30)		
Non-cognitive problems score		-1.66*		-0.41*
		(0.30)		(0.18)
Literacy score				0.73*
				(0.23)
Child obs.	3458	3107	3567	2837

Table H.3 shows the results of the OLS regression for the indicator equal to one if the mother expects the child to obtain a college degree in the future. Columns (1) and (2) show that there is a statistically significant gap for children at the start of formal schooling at ages 6-7, which persists even after controlling for teacher-reported literacy and mother-reported non-cognitive problems. Columns (3) and (4) show that this female-favourable gap in parental expectations exists for older children ages 14 to 15, around which time college preparation decisions are made. Therefore, it is possible that parents might have higher expectations for the educational attainment of their girls, which underscores the importance of literacy development in their daughters.

Table H.3: Gender gap in the expected probability that the child will obtain a college degree

	Age 6-7		Age 14-15	
	(1)	(2)	(3)	(4)
Child female	0.14*	0.09*	0.16*	0.07*
	(0.01)	(0.02)	(0.02)	(0.02)
Family income	0.05*	0.03*	0.05*	0.04*
	(0.01)	(0.01)	(0.01)	(0.01)
Mom coll+	0.21*	0.16*	0.18*	0.12*
	(0.02)	(0.02)	(0.02)	(0.02)
Num. sib.	-0.04*	-0.04*	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Two parents	0.07	0.03	0.14*	0.08*
	(0.04)	(0.04)	(0.03)	(0.03)
Literacy score		0.14*		0.19*
		(0.01)		(0.01)
Non-cognitive problems		-0.07*		-0.07*
score		(0.01)		(0.01)
Child obs.	3455	2766	2445	1929

I Gender gap in income processes

Tables I.4 and I.5 present the estimates of wage and non-labour income process with an additional control of child's gender. Child gender is not a significant predictor of hourly wages or the available non-labour income.

Table I.4: Wage process estimates controlling for child's gender

Log hourly wage mother		
Child age	0.000	-0.000
	(0.006)	(0.006)
Mother age	0.049**	0.049**
	(0.015)	(0.015)
Mother age sq.	-0.000*	-0.000*
	(0.000)	(0.000)
Mother coll+	0.332***	0.332***
	(0.017)	(0.017)
Child female		-0.010
		(0.015)
Const.	1.591***	1.601***
	(0.329)	(0.330)
Mother employed		
Available non-labour income	-0.000***	-0.000***
	(0.000)	(0.000)
Num. sib.	-0.288***	-0.288***
	(0.021)	(0.021)
Child female	0.065	0.063
	(0.043)	(0.044)
Child age	0.072***	0.072***
	(0.018)	(0.018)
Mother age	0.384***	0.384***
	(0.036)	(0.036)
Mother age sq.	-0.004***	-0.004***
	(0.000)	(0.000)
Mother coll+	0.414***	0.414***
	(0.048)	(0.048)
Const.	-7.376***	-7.374***
	(0.767)	(0.767)
Observations	8439	8439

Table I.5: Non-labour income process

	Single		Two-parent	
	(1)	(2)	(3)	(4)
Child age	-37.70***	-37.78***	-24.38	-25.08
	(8.25)	(8.24)	(16.55)	(16.56)
Mother age	-8.85	-9.00	65.81	67.65
	(7.67)	(7.66)	(34.98)	(35.00)
Mother age sq.	0.11	0.12	-0.57	-0.59
	(0.09)	(0.09)	(0.40)	(0.40)
Mother coll+	-54.00***	-54.72***	408.48***	409.86***
	(15.66)	(15.65)	(29.06)	(29.07)
Num. sib.	64.70***	64.53***	74.66***	74.82***
	(6.14)	(6.13)	(14.37)	(14.37)
Child female		-20.35		-44.26
		(14.16)		(28.40)
Const.	419.11*	431.66**	-593.50	-612.44
	(164.22)	(164.32)	(750.84)	(750.81)
SD shock	289.23***	288.99***	1160.62***	1160.42***
	(5.84)	(5.83)	(10.31)	(10.31)
N	1804	1804	6759	6759