IFS

Institute for Fiscal Studies

22/54

Working paper

Michele Giannola

# Parental investments and intra-household inequality in child human capital: evidence from a survey experiment

Economic and Social Research Council

# Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Survey Experiment

Michele Giannola*

First Draft: April 2020
This version: December 2022

**Abstract**

Intra-household inequality explains up to 50 percent of the cross-sectional variation in child human capital in the developing world. I study the role played by parents' educational investment to explain this inequality and its determinants. To mitigate the identification problem posed by observational data, I design a survey experiment with poor households in India. I develop new theory-driven survey measures based on hypothetical scenarios that allow me to separately identify parents' beliefs about the human capital production function *and* their preferences for inequality in children's outcomes, as well as study the role of household resources. I find that investment decisions are driven by efficiency considerations rather than inequality concerns over children's final outcomes. Because parents perceive investment to be 12 percent more productive for the higher-ability child, they allocate 10 percent more educational inputs to this child. Resources are important, as constrained parents select more unequal allocations. Counterfactual simulations indicate that policy interventions can have important intra-household distributional impacts through parents' behavioural responses.

*University of Naples Federico II, CSEF and the IFS. Address: Via Vicinale Cupa Cintia, 21, 80126 Naples, Italy. Email: michele.giannola@unina.it. I am very grateful to Orazio Attanasio and Imran Rasul for their support at each stage of this project, and to Alison Andrew, Richard Audoly, Samuel Berlinski, Richard Blundell, Teodora Boneva, Anne Brockmeyer, Gabriella Conti, Janet Currie, Mimosa Distefano, Carlo Galli, Costas Meghir, Rohini Pande, Francesca Salvati, Michela Tincani, Marcos Vera-Hernandez, and many seminar participants at Bank of Italy, Bank of Spain, Stockholm University, University of Bristol, University of Naples Federico II, UC3M, UCL, the IFS, Yale, ESPE 2022 and SIEP 2022 for helpful comments and discussions. I would also like to thank Lina Cardona-Sosa for her invaluable help with the fieldwork. All errors are my own. The study has obtained ethical approval from the UCL Research Ethics Committee (IRB Approval Number 16727/001) and the registry trial can be found at https://www.socialscienceregistry.org/trials/5444. This paper was previously circulated as "Parental Investments, and Intra-household Inequality in Child Human Capital: Evidence from a Lab-in-the-field Experiment".

# 1 Introduction

Intra-household inequality is key for the measurement and understanding of poverty and inequality (Haddad & Kanbur (1990); Chiappori & Meghir (2015)), and has important consequences for the effective targeting of social programmes in the developing world (Brown, Ravallion, & Van De Walle (2017)). Children are the most vulnerable to intra-household inequalities (Dunbar, Lewbel, & Pendakur (2013)). Because the early years are fundamental to the process of human capital formation (Currie & Almond (2011); Brito & Noble (2014); Heckman & Mosso (2014)), intra-household allocations can have profound consequence for wellbeing over the life-cycle.

This paper focuses on intra-household inequality in children's human capital outcomes, and studies the role played by parents' educational investment to explain this inequality. Do parents invest more in higher-ability children, exacerbating inequality in a society, or do they compensate for endowment differences, acting as an equalising agent? What are the determinants of this decision? Do parents perceive some children to have higher returns to education than others? Do they care about inequality in children's final outcomes? Are these decisions affected by household resources? Given that the effects of public programmes targeting children are mediated by parents' behavioural responses, answers to these questions are fundamental for the design of policies aimed at improving child wellbeing and reducing inequalities. This is particularly true in developing countries, where social protection systems are less well established and families are the primary providers of material support to their children.

The empirical analysis of whether parents' investment reinforces or compensates initial differences is plagued by the fact that measures of child endowments, at birth or early in life, include both a genetic component *and* a behavioural component of parental nurturing (Rosenzweig & Wolpin (1988)).[1] In this case, a positive relation between early levels of human capital and the subsequent demand for investment inputs is spurious as it would reflect the correlation in parents' behaviour over time. A preference for one specific child, for example, would result in a higher endowment for the preferred child *and* higher investment in that same child.[2]

Parents' behaviour is informed by their preferences, resources and perceptions of the process of human capital formation. However, identifying the separate role played by these factors in the reduced-form demand equations for child human capital investment is also challenging. The reason for this challenge is a twofold identification issue. First, input choices are consistent with many alternative specifications of preferences and beliefs (Manski (2004)), which are typically unobserved in survey data. Second, resource constraints might limit parents' ability to invest (Lochner & Monge-Naranjo (2012)), breaking the link between observed outcomes and preferences. Nonetheless, understanding what determines these demand relationships is important when thinking about many polices, and relevant for welfare analysis (Caplin & Martin (2021)). Consider the case of information interventions for parents, a commonly proposed policy to increase parental investment and improve child outcomes: changing parents' perceptions will be effective insofar as demand is driven by beliefs rather than preferences. Similarly, the effects of programmes that transfer educational resources to poor households

[1] For example, a large literature documents the important of maternal behaviour during pregnancy on child health at birth (e.g., Rosenzweig & Schultz (1983)). A similarly large literature demonstrates that child human capital early in life is endogenously determined by parents' early investment (see Currie & Almond (2011) for a review).

[2] Aizer & Cunha (2012) report a positive correlation between parents' pre- and post-natal human capital investment.

might depend on how parents allocate these resources between siblings. For some specifications of parents' preferences, this policy could increase cross-sibling inequality (Barrera-Osorio, Bertrand, Linden, & Perez-Calle (2011)).

This paper reports the results of a survey experiment purposely designed to study parents' human capital investment in their children, that I conduct with poor households in urban Odisha, India. I develop theory-driven survey measures that allow me to identify the causal effect of child human capital endowment on parents' educational investment. The strategically-designed survey instruments further allow me to separately identify the primitive parameters in the reduced-form demand functions: parents' beliefs about the human capital production function *and* their preferences for inequality in children's outcomes, as well as study the role of household resources. I complement these data with information on *actual* parental investment to validate the experimental strategy.

The survey experiment consists of two stages. In the first stage, I identify parents' beliefs about the human capital production function. The approach used to elicit these beliefs builds on the seminal work by Cunha, Elo, & Culhane (2013), and consists in presenting a series of hypothetical situations (scenarios) to the respondent and elicit information on individual expected outcomes. By varying the characteristics of the scenarios one at a time while keeping *all* other characteristics of the environment constant, I identify the perceived returns to specific inputs that enter the child human capital production function and are relevant for the investment decision. Specifically, respondents are asked to report what they except the child outcome would be in each of a set of scenarios that vary in terms of child baseline ability and parents' educational investment. Comparing answers between scenarios, I identify the perceived returns to these inputs, as well as their perceived complementarity or substitutability.

Having identified beliefs, I then collect parents' stated investment choices. In this stage of the experiment, respondents are presented with hypothetical scenarios describing a family that makes an investment decision regarding their children's education, and asked to choose their preferred allocation of resources between two children with varying baseline ability. The experiment introduces exogenous variation in child endowments that I use to identify the reduced-form equations relating child ability to parents' investment. I then combine choices made in the second stage experiment with beliefs, to identify parents' preferences free from other confounding factors. Importantly, by directly eliciting information about the *perceived* production function, I identify preferences without imposing that parents' beliefs about the process of human capital accumulation correspond to the true process; an assumption that does not hold in practice (Cunha, Elo, & Culhane (2013); Boneva & Rauh (2016)), but upon which earlier work relies (e.g., Behrman, Pollak, & Taubman (1982)).

Several key results emerge from this study. First, I find that parents perceive child baseline ability and educational investment to be productive in the human capital production function. A one-standard-deviation increase in baseline ability is perceived to increase earnings at age 30 by 15 percent; a similar increase in investment is expected to increase earnings by 28 percent. Moreover, parents perceive the two inputs to be *complements*: They believe that investment is 12.6 percent more productive for higher-ability children. This perceived complementarity generates an incentive for parents to reinforce initial differences between children *if* they seek to maximize the returns from their investment.

Experimental results from the second-stage experiment reveal that parents reinforce baseline differences between children. Specifically, I show that a one-standard-deviation increase in the difference in children's abilities leads to a 9.7 percent increase in the share of resources allocated to the higher-achieving child. I also find that reinforcement is significantly weaker when household resources are higher, with the lower-achieving child penalty in investment being roughly halved. Choices made in the second-stage experiment, combined with elicited beliefs from the first stage, imply that parents have a low aversion to inequality over their children's human capital outcomes. Although I reject that parents are pure returns-maximizers, the results suggest that, in this setting, investment choices are primarily driven by efficiency considerations rather than inequality concerns over final child outcomes.

By using detailed information on parents' educational investments collected separately for each of their *own* children (in terms of both monetary and time investment), I then show that there is a robust relation between parameters identified in the experiment and *actual* investments. Respondents identified as less inequality averse spend more unequally in their own children's education, favouring their higher-ability child, with the effects primarily coming from private tuition expenditure. I also find that, while parents spend a similar amount of time playing with their children, they spend more time with their higher-ability child on school-related activities. Although I do not attach a causal interpretation to these patterns, they are reassuring and add credibility to the research design. They also suggest that investment behaviour observed in the experiment might extend beyond educational expenditure to other high-stakes investment choices that have important long-term consequences for child wellbeing. These results confirm recent evidence showing that the two approaches of using stated or actual choices yield similar results when the hypothetical scenarios are realistic and relevant for the respondent (Mas & Pallais (2017); Wiswall & Zafar (2018)).

Finally, I use the experimental results to simulate the effects of a policy that increases parents' beliefs about the productivity of investment, while holding fixed the educational budget. This policy resembles some of the existing programmes serving economically disadvantaged families that aim at changing parental behaviour by affecting information about investments in children (e.g., York, Loeb, & Doss (2019); List, Pernaudet, & Suskind (2021)). I study the effects of this policy on parents' allocation of investments and on children's human capital. I show that while this policy is welfare improving for the household as a whole, it has uneven effects on individual children's wellbeing because of parents' endogenous responses. This counterfactual experiment highlights that, to the extent that families are the ultimate decision makers, it is necessary to consider behavioural responses to understand the impacts of policies on the wellbeing of individual household members. It further suggests the need for future research to explore the intra-household distributional impacts of these interventions.

This paper makes several novel contributions to the literature. First, it contributes to the literature focusing on the role of subjective beliefs as a determinant of parents' human capital investment. This literature has documented the existence of information frictions such that parents underestimate the returns from their investment (Cunha, Elo, & Culhane (2013); Boneva & Rauh (2016)). These frictions are typically more pronounced for poor parents (List, Pernaudet, & Suskind (2021)), and have been related to inequalities in human capital investments *between* families (Boneva & Rauh (2018); Attanasio, Cunha, & Jervis (2019); Bhalotra, Delavande, Gilabert, & Maselko (2020); Attanasio, Boneva, & Rauh (2022)). There is also evidence that beliefs are malleable (List, Pernaudet, & Suskind (2021)), and correcting these biases has a causal impact on parents' investment (Dizon-Ross (2019)).[3] This work contributes to this literature by showing that parents' beliefs about the human capital production function matter to explain differences in investment between children *within* the same family, beside their importance to explain inequalities *between* families. This study is also one of the first to analyse the role of parents' perceived returns to investment for school-aged children in a developing country.[4] From a methodological perspective, in contrast with most previous studies, I use information on the perceived production function to further *"separate production technology from parental preference"*, as recently invoked by Yi, Heckman, Zhang, & Conti (2015).

Second, the paper provides novel experimental evidence on parents' allocations of human capital investment between children with varying endowments. Because variation in child characteristics is non-random, this literature has used family fixed-effects models or natural experiments, in an attempt to overcome this identification challenge (see Almond & Mazumder (2013) for a review). In this paper, I exploit the experimentally induced variation in child endowments for identification. This approach has several advantages. First, it limits the role of unobserved preferences over specific child attributes, such as gender (Barcellos, Carvalho, & Lleras-Muney (2014)) or birth order (Jayachandran & Pande (2017)), in driving parents' investment. By design, the use of hypothetical scenarios fixes these child characteristics, as well as other features of the environment that might be unobserved to the researcher. Second, unlike observational data, experimental allocations are both *private*, as they cannot be shared among children, and *assignable*, as they are specific to one individual child known to the researcher.

Moving beyond the primitive question of whether parents reinforce of compensate initial differences, this paper relates to the literature studying parents' preferences for investing in their children (Becker & Tomes (1976); Griliches (1979); Behrman, Pollak, & Taubman (1982); Behrman (1988); Pitt, Rosenzweig, & Hassan (1990)). While in this literature identification is achieved under the assumption that the *perceived* production function corresponds to the *actual* one, the experimental approach allows me to identify parental preferences under far weaker assumptions.[5]

Third, by showing that household resources have important implications for the allocation of human capital investments between children *within* the same family, this paper complements the literature

[3]Beyond affecting parents' investment in education, providing information to individuals has been shown to affect decision-making across different domains (e.g., Jensen (2010), Dupas (2011), Liebman & Luttmer (2015), Fitzsimons, Malde, Mesnard, & Vera-Hernández (2016), Hjort, Moreira, Rao, & Santini (2021))

[4]Attanasio, Cunha, & Jervis (2019) and Bhalotra, Delavande, Gilabert, & Maselko (2020) consider subjective beliefs for much younger childre in Colombia and Pakistan, respectively.

[5]In a recent paper, developed independently from this study, Berry, Dizon-Ross, & Jagnani (2020) use a set of experiments to study parental preferences. In the experiments, parents allocate lottery tickets for a one-hour tutoring on a test that is provided to the children immediately after the experiment and payments are made to each child based on their performance on the test. The identification strategy in Berry, Dizon-Ross, & Jagnani (2020) relies on varying the function that maps parents' choices to child outcomes by shocking the *short-run* returns to invest in different children (e.g., by rewarding more the test scores of one child over another). The identification strategy that I use does not exploit exogenous changes to production function, but uses direct information on the perceived production function. Because I do not alter the function mapping inputs to outcomes, I identify preferences conditional on parents' *own* beliefs, without changing the environment in which parents typically make their choices. This further allows me to study the role of both parents' preferences *and* beliefs, which have never been jointly analyzed. One advantage of the approach used by Berry, Dizon-Ross, & Jagnani (2020) is that choices made by parents in the experiment are incentivised by delivering real monetary payments, while I rely on hypothetical choices. Reassuringly, Berry, Dizon-Ross, & Jagnani (2020) replicate their main findings in a hypothetical survey experiment. This squares well with recent evidence pointing to the fact that the two approaches yield similar results (Mas & Pallais (2017); Wiswall & Zafar (2018)).

investigating the role that financial constraints have in explaining gaps in educational investments *between* low- and high-SES families (Lochner & Monge-Naranjo (2012); Kaufmann (2014); Solis (2017)). My contribution to this literature is to show that the lack of resources might prevent parents from adequately investing in *all* their children, leading them to select their higher-achieving child, and leaving more vulnerable children at a considerable risk. This result is consistent with early findings in Behrman (1988), showing that parents favour better endowed children in the lean season, and more generally with the idea that "*discrimination is stronger in a time of crisis*" (Duflo (2005)). In turn, this suggests that reducing poverty could disproportionately benefit weaker children.

This paper also contributes to a growing body of evidence pointing at the importance of considering intra-household inequality to understand differences across individuals in a society.[6] While this literature focuses on inequalities between different groups of individuals living in the same household (e.g., men *vs.* women; adults *vs.* children), I document the importance of intra-household inequality *between* children. Moreover, the results from the counterfactual experiment highlight the importance of considering parents' behavioural responses to predict the effects of polices on individual welfare.

Finally, in terms of field methodology, this paper relates to a growing literature that uses strategically-designed survey measures to collect data on individual beliefs, and elicit stated choices to understand behaviour and identify primitive parameters of interest (Caplin (2016, 2021); Attanasio (2021); Stantcheva (2022)). The method that I use to elicit parents' beliefs is not novel (e.g., Cunha, Elo, & Culhane (2013), Boneva & Rauh (2018), Attanasio, Boneva, & Rauh (2022)). But distinctly from previous work, that typically only correlates elicited beliefs with observed choices, I further use these data to identify preferences. Two recent papers by Delavande & Zafar (2019) and Adams-Prassl & Andrew (2020) also combine information on beliefs and hypothetical choices collected with strategically-designed survey measures to understand the determinants of individual behaviour in relation to university and marriage decisions. The present paper is closest to these in terms of field methodology.

The paper is organized as follows. The following section presents two stylized facts that motivate this study. Section 3 presents the conceptual framework that I use to study intra-household inequality in children's human capital outcomes, and clarifies the identification challenges posed by observational data. Section 4 presents the experiment and how it solves the identification challenges. Section 5 describes the setting and the data. The results are presented in section 6. Section 7 uses a counterfactual exercise to discuss the implications of the results for policy and welfare, and Section 8 concludes.

## 2 Motivating Evidence

Two empirical facts motivate this study. The first is presented in Figure 1, which plots the share of total variation in child human capital (as measured by educational attainment) that can be attributed to within-household and between-households variation across a number of developing countries. To perform this decomposition, I use the Mean Log Deviation (MLD) measure of inequality (Ravallion (2015)), which can be exactly separated into a within-group component and a between-groups component (see Online Appendix A). The figure shows that, for a large set of developing countries, intra-household

[6]See Haddad & Kanbur (1990); Lise & Seitz (2011); Dunbar, Lewbel, & Pendakur (2013); Chiappori & Meghir (2015); Brown, Ravallion, & Van De Walle (2017); Brown, Calvi, & Penglase (2020); Calvi (2020).

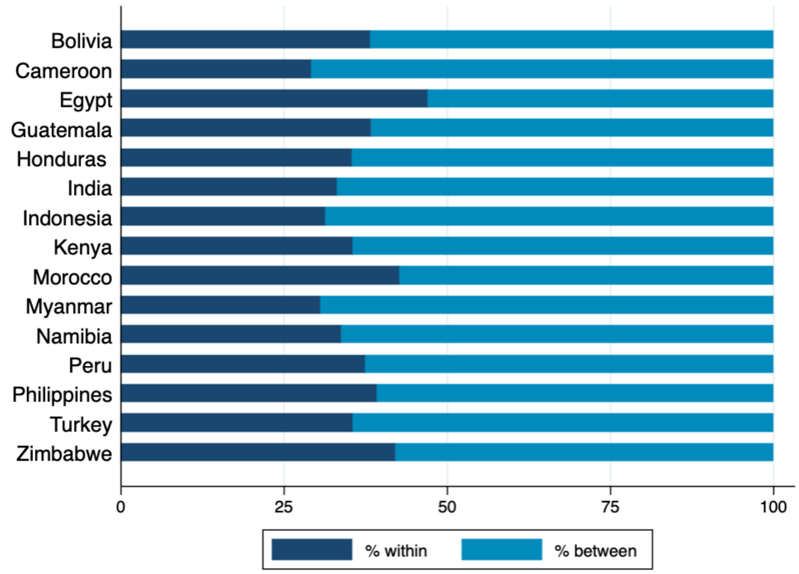


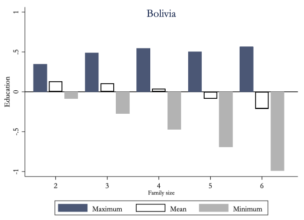

Figure 1: Contribution of Intra-household Inequality to Child Human Capital Inequality

*Notes:* This figure plots the within-household and between-households component of the Mean Log Deviation (MLD) measure of inequality. The outcome variable is educational attainment. I use an age-standardized z-score, where the reference group consists of children in the same country and birth cohort. Source: Development and Health Survey (DHS) and NFHS for India.

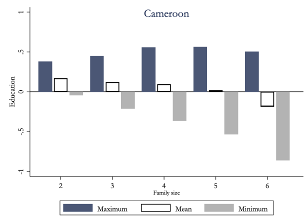
inequality explains between 30 and 50 percent of the cross-sectional variation in child human capital. In India, the country under study in this paper, inequality between siblings amounts to 33 percent of the overall inequality in educational attainment (Appendix Figure A.1 reports similar results for age-standardized test scores).

Figure 2 presents the second empirical fact motivating this paper. The figure shows how the distribution of child human capital within the family varies with household size. In particular, the figure focuses on the mean (in white), the maximum (blue), and the minimum (grey) of this distribution, that is the level of human capital of the *highest* and *lowest-achieving* child, and the *average* level of human capital in the household. This relation is plotted for the same set of developing countries in Figure 1, including India (Panel F) and the state of Odisha, the setting for this study (Panel G).

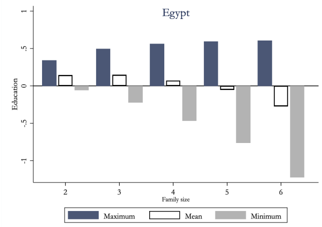
The figure reveals several patterns that are strikingly similar across countries. First, there is a negative relation between average child outcomes and family size. This is a relatively well documented fact that dates back to Gary Becker's Quality-Quantity (Q-Q) model, and can be explained by the fact that in larger families there are less per-capita resources, so that on average each child receives less human capital investments (Becker & Lewis (1973)). Second, Figure 2 shows an interesting and less documented relation between the distribution of human distribution in the family and household size. Notably, while the human capital of the highest-achieving child in the family does not vary with family size, the outcome of the least successful child steeply declines as household size increases. This relation is robust to a series of robustness checks (that I report in Online Appendix B), including controlling for child gender and birth order, as well as considering the endogenous selections of families into different



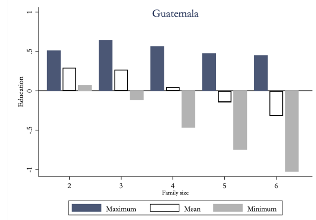

(A) Bolivia



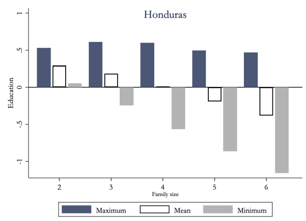

(B) Cameroon



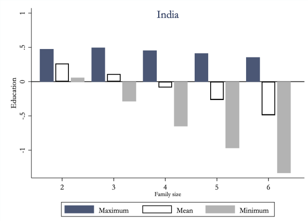

(C) Egypt



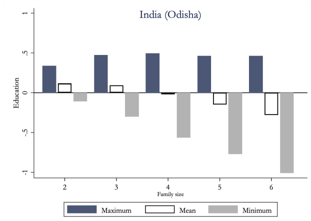

(D) Guatemala



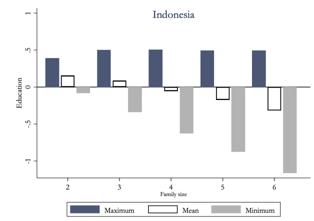

(E) Honduras



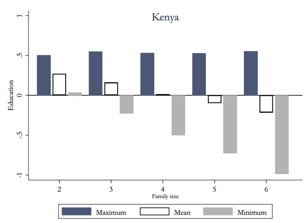

(F) India (whole)



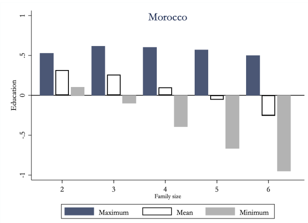

(G) Odisha (India)



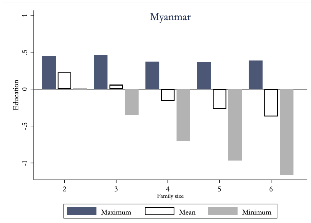

(H) Indonesia



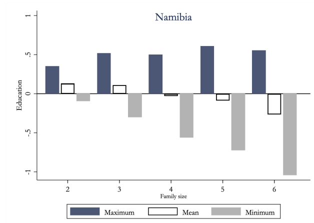

(I) Kenya



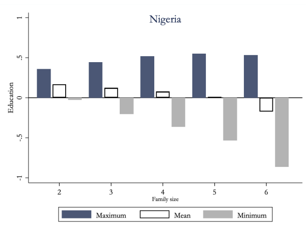

(J) Morocco



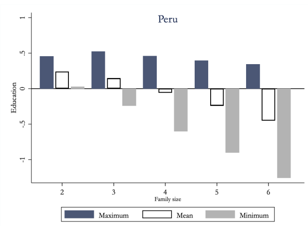

(K) Myanmar



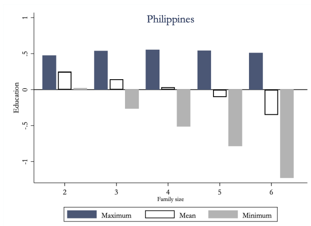

(L) Namibia



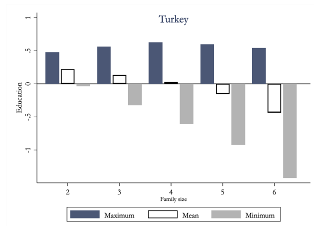

(M) Nigeria



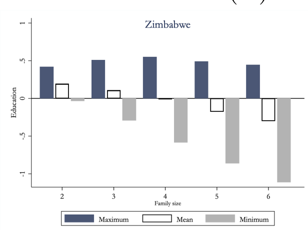

(N) Peru




(O) Philippines




(P) Turkey




(Q) Zimbabwe

Figure 2: Human Capital Distribution by Family Size

*Notes:* The figure plots the relationship between family size (x-axis) and the mean (in white), the maximum (in blue) and the minimum (in grey) levels of human capital within the household (y-axis). This figure is constructed as follows. For each family in the sample, I compute the maximum, minimum and mean levels of human capital achieved by children in that family. For each family size, I then average across families. The outcome variable is educational attainment. I use an age-standardized z-score, where the reference group consists of children in the same country and birth cohort. Thus coefficients are expressed in standard deviations units. Source: Development and Health Surveys (DHS) and NFHS for India.

levels of fertility. Therefore, the negative correlation between average child quality and quantity can be explained by reductions at the bottom of the human capital distribution.[7]

The figure paints a more nuanced picture of the relation between child human capital and family size. While the Q-Q model implicitly assumes that parents invest similarly in *all* children, these patterns are consistent with an unequal allocation of human capital investment between children in the family. This suggests the possibility of a *behavioural* origin underlying intra-household inequality in child outcomes: unequal parents' human capital invetment. This investment strategy might be particularly detrimental for the human capital of children in larger families, because of less per-capita resources. Unequal investments could explain at the same time the steep decline in the *minimum*, the flat gradient in the *maximum*, and the reduction in *average* level human capital reported in Figure 2. To understand the role played by intra-household allocations for inequality in child human capital outcomes, this paper analyses whether parents make unequal educational investments between children, and investigates the determinants of parental behaviour, studying the role of preferences, beliefs and constraints.

# 3 Conceptual Framework

This section develops a simple theoretical framework to study how parents allocate resources between children, and highlights the challenges posed by observational data. I use this framework to inform the design of the survey instruments used in the field, guide the empirical analysis, and interpret the findings.

## 3.1 Preferences and constraints

As in standard models of human capital investments, parents derive utility from their children's human capital outcomes. The dependence of the utility function on child human capital can be motivated by parents' altruism towards the children, or by the fact that children might represent a source of support during old age (the present study does not attempt to separate these alternative motives). To focus on the allocation of resources among children, I further impose a standard separability assumption between parents' consumption and child outcomes (Behrman (1997)). Preferences over child outcomes are specified as a Constant Elasticity of Substitution (CES) utility function and expressed as:

$$U(H_1 ... H_n) = (c_1 H_1^{\rho} + c_2 H_2^{\rho} + ... + c_n H_n^{\rho})^{\frac{1}{\rho}} \tag{1}$$

where $H_i$ is child $i$ human capital (typically her adult earnings or educational attainment), $c_i$ is a child-specific preference that might depend on the child's characteristics (e.g., a preference for sons over daughters), and $\rho$ is the preference parameter that regulates parents' aversion to inequality in child outcomes. The functional form assumption is standard in the literature on intra-household allocation of resources (e.g., Behrman, Pollak, & Taubman (1982) and Behrman (1988)). The CES specification is very flexible in that it allows a complete range of productivity-equity trade-offs: At one extreme when $\rho = 1$, the indifference curves are linear and there are no inequality concerns. In this case parents act

[7] Interestingly, Aizer & Cunha (2012) identify a similar pattern in a sample of poor households in the US.

as returns-maximizing agents. The opposite case is the Rawlsian case when $\rho \to -\infty$: utility curves are L-shaped and parents act to equalize child outcomes. In between these two cases, parents trade-off efficiency and equity concerns. Parents choose educational investment in their children $I_i$ to maximize their utility subject to two constraints. The first is a budget constraint. As this is a static one-period model without saving or borrowing, this is expressed as:

$$y = I_1 + I_2 + ... + I_n \tag{2}$$

where $y$ is the total educational budget, and the price of investment is normalized to one. One can imagine a two-stages budgeting process: in the first stage parents decide the amount of resources to spend on their children's education, and then how to distribute these resources between children.[8] The second constraint faced by the family is a human capital production function, which maps inputs into later life outcomes. This is given by:

$$H_i = f(A_i; I_i; Z_i) \tag{3}$$

where $A_i$ is child baseline ability (her endowment), $I_i$ are educational resources devoted to the child by her parents, and $Z_i$ are other child or family characteristics.[9,10]

## 3.2 Subjective beliefs

Models of intra-household allocations of resources rely on strong assumptions about parents' knowledge of the human capital production function, as they assume that parents have perfect information about the "true" technology of skill formation in (3). This assumption is a strong one, and has been shown not to hold in a variety of contexts. For example, Cunha, Elo, & Culhane (2013), Boneva & Rauh (2016) and Attanasio, Cunha, & Jervis (2019) show that parents hold inaccurate beliefs about the productivity of different inputs entering the human capital production function, and typically underestimate such returns. To incorporate these information frictions into the model, I introduce the *perceived* human capital production function, which is specified as:

$$H_i = \tilde{f}(A_i; I_i; Z_i) \tag{4}$$

[8] As in a standard Q-Q model, family size $n$ does not have a direct effect on household resources $y$, but reduces the amount of per-capita resources available $y/n$.

[9] The model and empirical analysis could be extended to other dimensions of child endowments such as physical health; what matters is that endowments affect the returns to investment, and are relevant to determine child outcomes.

[10] The model relies on several additional assumptions that are standard in the literature on parents' intra-household allocations. First, I assume that the total number of children in the family is exogenous. Second, I consider an unitary model of the family where parents act as a single optimizing agent. Third, I define inequality aversion over human capital outcomes rather than over consumption. If parents are inequality averse over consumption, they could maximize returns at the investment stage and redistribute consumption later with transfers. However, for poor families as the ones considered in this paper large monetary transfers from parents to children are unlikely to take places later in life. Relatedly, there is a literature looking at parental bequests, suggesting that parents do not equalize outcomes ex post but bequest similarly all children (Behrman & Rosenzweig (2004)). I discuss some of these assumptions and how they can be relaxed in greater detail in Appendix B.1. For example, I consider an extension of the model that accounts for endogenous family size, and show that in the model fertility choices also depend on parents' preferences.

This is allowed to differ from the actual human capital production function so that $f \neq \tilde{f}$, capturing the fact that parents have incomplete information about how inputs map into future child outcomes. Equations (3) and (4) play two very different roles: The former describes the actual process of child development, while the latter describes parents' subjective beliefs about this process, and is the relevant constraint taken into account when deciding children's human capital investments.[11] The maximization problem of the parents results in a policy function $I_i^*$ describing the optimal level of human capital investment for each child. This is a function of parents' resources and preferences, the perceived production function, and child baseline ability.

### 3.3 Identification

To illustrate the key challenges to identify the demand function $I_i^*$ and its behavioural determinants using observational data, I put some more structure on the problem by specifying a functional form assumption for the production function. I assume that this is Cobb-Douglas, and express it as:

$$H_i = A_i^{\alpha} I_i^{\beta} \tag{5}$$

where $\alpha$ and $\beta$ are the returns to ability and investment. I assume that the perceived technology is also Cobb-Douglas, but parents do not know the true productivity parameters, and specify $\tilde{f}$ as:

$$H_i = A_i^{a} I_i^{b} \tag{6}$$

where $a$ and $b$ are the *perceived* returns, and these are allowed to differ from *actual* returns.[12]

Solving for the optimal level of investment in each child (see Appendix B.2), I derive the following equilibrium allocation rule:

$$\log\left(\frac{I_i^*}{I_j^*}\right) = \frac{a\rho}{1-b\rho}\log\left(\frac{A_i}{A_j}\right) + \frac{1}{1-b\rho}\log\left(\frac{c_i}{c_j}\right) \tag{7}$$

This is the structural relation of interest, relating the allocation of investments to parents' preferences ($\rho$, $c_i$ and $c_j$), their beliefs ($a$ and $b$), and children's baseline abilities ($A_i$ and $A_j$). The primitive parameters in (7) are not identifiable using survey data, as these typically only include information on

[11] I do not consider how parents form these beliefs and whether these evolve over time. There are both theoretical and empirical reasons for doing so. First, the model is static so what matters to determine choices is the beliefs that parents hold at a particular point in time. Second, the data that I use are not longitudinal in nature, making them not appropriate to answer these questions. A literature in psychology suggests that individuals use heuristics to form expectation (Tversky & Kahneman (1974)). A small body of work in economics has looked at how individual form beliefs and how these evolve (Di Tella, Galiani, & Schargrodsky (2007)). The study of how parents form beliefs about the process of child development, and how these beliefs change over time should be the focus of future research.

[12] While a more flexible specification for the production technology could have been used – for instance one that allows richer patterns of substitutability between inputs – previous research has found the Cobb-Douglas to be a reasonable specification in the Indian setting (Attanasio, Meghir, & Nix (2020), see also Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020); and Attanasio, Bernal, Giannola, & Nores (2020) for examples from another developing country). Attanasio, Cunha, & Jervis (2019) find that this functional form well approximates the production function as perceived by the parents. Importantly, Cobb-Douglas production allows to derive closed form solutions for parents' investment and to point identify parents' inequality aversion, a key parameter in this literature (see Section 4).

investments and some imperfect proxies of children's endowments, but not on preferences or beliefs.[13] The corresponding reduced-form equation that can be estimated from observational data is:

$$\log\left(\frac{I_i^*}{I_j^*}\right) = \gamma \log\left(\frac{A_i}{A_j}\right) + \epsilon \tag{8}$$

Where $\epsilon$ is an error, including the unobserved (to the researcher) term depending on $c_i$ and $c_j$ in (7). I now describe two key challenges in the identification of equations (7) and (8).

**1. The reduced-form coefficient $\gamma$ is generally biased.** Identification of the reduced-form parameter $\gamma$ requires that differences in children's abilities are orthogonal to the error term $\epsilon$ in (8). This is usually not the case as measures of child endowment comprise an endogenous component of parental nurturing (Rosenzweig & Wolpin (1988)). In this case, a positive relation between baseline ability and subsequent investment would be spurious as it would reflect the correlation in parents' behaviour over time, as reported for example by Aizer & Cunha (2012).[14] A preference for one specific child, for example, would result in a higher endowment for the preferred child *and* higher investment in that same child, implying that the reduced-form coefficient $\gamma$ is upward biased. In general, a regression of parents' investment on children's ability does not identify the reduced-from coefficient $\gamma$, because the error term is not orthogonal to the difference in children's abilities in (8).

**2. The primitive parameters in the demand function can not be recovered from observed choices.** Equation (7) shows that the interplay between preferences and perceived returns determines the allocation of investment between children. Without additional information about the perceived production function it is not possible to derive conclusions about parental preferences using observed choices, as these are consistent with many alternative specifications of preferences and beliefs. Without imposing strong assumption on beliefs – such that the parameters of the perceived production function correspond to the parameters of the actual technology – one can not identify parents' preferences. For example, finding that investments do not vary with child ability could either mean that parents are averse to inequality in child outcomes, or that they perceive the returns to ability to be low: in both cases the corresponding reduced-form parameter is close to zero.

As detailed in the next section, to overcome these identification challenges I use strategically-designed survey questions (Ameriks, Briggs, Caplin, Lee, et al. (2020)) that allow me to identify both the reduced-form relation between child ability and parents' investment *and* the primitive parameters in the structural demand function. The survey instruments that I develop are explicitly informed by the theoretical framework, and designed to bypass the shortcomings of observational data. These novel data are collected in conjunction with data on *actual* parental behaviour that I use to validate the experimental results.

[13]Previous studies have used weight at birth (Datar, Kilburn, & Loughran (2010)), health status (Leight (2017)), or cognitive abilities (Adhvaryu & Nyshadham (2016)) as proxies of endowments.

[14]Aizer & Cunha (2012) report that child pre- and post-natal investment are correlated, also when maternal fixed effects are included.

# 4 Survey Experiment

The conceptual framework illustrates the challenges that observational data pose for the identification of key relations and parameters of interest. To overcome these identification challenges, I design and implement a survey experiment with parents of primary school children in urban Odisha, India. In the experiment, I use novel strategically-designed survey measures guided by the theoretical framework. This section describes the experimental measures and procedures used in the field, and how I combine different measures to identify primitive parameters of interest. The experiment consists of two stages. In the first-stage experiment, that I describe in section 4.1, I elicit parents' beliefs about the human capital production function. In the second-stage, described in Section 4.2, I collect parents' stated investment choices. I then combine beliefs and choices to back-out parents' preferences for inequality in child outcomes, using the procedure detailed in Section 4.3.

## 4.1 First-stage experiment: Beliefs

**Measurement.** To elicit parents' beliefs about the human capital production function, I build on the seminal work by Cunha, Elo, & Culhane (2013), and use hypothetical scenarios.[15] This strategy consists in presenting a series of hypothetical situations to the respondent and elicit information on individual expected outcomes. The experimental procedures used in the field worked as follows. Surveyors presented to the respondent a series of hypothetical stories (scenarios) about a representative family living in a neighbourhood like their own. The family was described as having *two* children and making decisions regarding their education. Guided by the theoretical framework, I focus on the role of perceived returns to child baseline ability ($A_i$), parental investment ($I_i$), and on their perceived complementarity or substitutability. To identify these perceived returns, I exogenously varied these characteristics between scenarios and asked the respondent what they expected the earnings of the child would be at age 30 (this corresponds to $H_i$ in the theoretical framework).

Specifically, the two children in the scenarios were described as attending the same primary school and identical in many aspects, but differing in *one* important characteristic: while one child – *Child H* – had an high baseline ability, the other child – *Child L* – had a low baseline ability. Specifically, to convey information about ability, *Child H* was described as being *"among the top three students in his/her class"*, while *Child L* was described as being *"among the bottom three students in his/her class"*. As it is common in many developing countries (e.g., Bhalotra, Delavande, Gilabert, & Maselko (2020)), parents in India value school performance and consider it an important indicator of their child ability. Previous research has demonstrated that parents base important investment decisions on their child academic performance (Dizon-Ross (2019)). This is also the first reliable and objective measure of child ability that parents have access to, and during piloting this description appeared clear and intuitive for the respondents.

Scenarios then varied the amount of monetary investment made by the family in the education of each child in terms of school fees, private tuition, stationary, books and other school related expenditures (these are all expenditure items that make part of the educational budget of households in the

[15] A similar empirical strategy has been used recently by Boneva & Rauh (2016), Boneva & Rauh (2018), Attanasio, Boneva, & Rauh (2022) and Conti, Giannola, & Toppeta (2022), among others.

experiment). Some scenarios described a *high* level of investment (identified as the 90th percentile of educational expenditure during piloting), while other scenarios described a *low* level of investment (corresponding to the 10th percentile of educational expenditure in the data). During the survey, enumerators explained (by explicitly stating) that these investments would help the children acquire new skills and progress through their educational careers, thus emphasising the *long-run* nature of the investments (Online Appendix C reports the exact wording used by enumerators during the experiment).

After presenting each scenario, respondents were asked to report what they expected the outcome would be for each child in terms of earnings at age 30. The respondent's answer was recorded, and the enumerator moved on to the next scenario. As is commonly done in developing countries (e.g., Delavande & Kohler (2009)), all hypothetical scenarios were presented with the help of visual aids that sketched the main features and made salient to the respondent the differences across scenarios (Appendix Figure A.2 presents an example of visual aid used in the field). As a robustness check and to gain a better understanding of the respondents' reasoning, parents were also asked to state what they believed the educational attainment of the child (measured in years of education) would be in each scenario. Finally, to understand whether *average* parental beliefs significantly differ by child gender, respondents were randomised in two groups: in one group the hypothetical scenarios described two children who were boys, while in the other group scenarios described one girl and one boy.

The use of hypothetical scenarios to identify parents' beliefs has several advantages. First, between hypothetical scenarios one can vary one input at the time while holding *all* other characteristics of the environment fixed, thus identifying the perceived productivity of that specific input. Second, the use of hypothetical scenarios allows me to identify beliefs without directly asking respondents about probabilities, which might be important in settings with low literacy levels as the one of this study (as described in Section 5). Compared to previous studies that describe different hypothetical families (e.g., Boneva & Rauh (2018)), presenting respondents with one single family with two children (rather than two distinct families with one child) has the additional advantage of holding fixed many characteristics of the environment that matter for child outcomes and vary between families (e.g., parental income and the family environment), but are unobserved to the researcher and might influence the responses to the hypothetical questions. The design absorbs these between-families differences, so only differences between the two children matter.[16]

To elicit subjective expectations using hypothetical scenarios one can either ask respondents about their own child or about a hypothetical one. Advantages and disadvantages of each method are discussed in Delavande (2014). I decided to ask parents about hypothetical children rather than their own as this allowed me to have control over *all* child characteristics. One key input of interest in this context is child baseline ability, and exogenous variation in ability would not have been possible if I asked respondents about their own children, as parents might base their answer to the hypothetical questions on some (unobserved) information specific to their own children.

**Identifying the perceived production function.** Comparing responses across scenarios and between children I identify: (i) the perceived returns to investment, (ii) the perceived returns to baseline

[16] It might still be possible that parents inferred from the description of the hypothetical scenarios that the higher-ability child enjoyed more studying compared to the lower ability-child and took this into account when answering.

ability, and (iii) the perceived complementarity or substitutability between these two inputs. For example, comparison of responses in the scenarios where investment is high to the corresponding scenarios where investment is low (holding fixed child ability) identifies the perceived returns to this input.

Formally, to characterise the perceived production function for child human capital I estimate the following empirical specification using ordinary least squares (OLS):

$$y_{i,j,k} = \alpha_0 + \alpha_1 A_k + \alpha_2 I_{j,k} + \alpha_3 A_k \times I_{j,k} + \eta_i + u_{i,j,k} \tag{9}$$

where $i$ indicates the respondent, $j$ the scenario and $k$ indicates one of the two children in each scenario. $y_{i,j,k}$ are expected (log) earnings, $A_k$ is a dummy variable that equals one if child $k$'s baseline ability is high, $I_{j,k}$ is a dummy that equals one if investment in child $k$ in scenario $j$ is high, and $\eta_i$ are respondent fixed effects. The coefficients $\alpha_1$ and $\alpha_2$ identify the perceived returns to baseline ability and investment, while the coefficient $\alpha_3$ identifies their perceived complementarity ($\alpha_3 > 0$) or substitutability ($\alpha_3 < 0$). Variants of this specification allow me to study whether *average* perceived returns vary by child gender, by comparing respondents randomised in one of the two groups described above.

### 4.2 Second-stage experiment: Investment choices

**Measurement.** In the second round of the experiment I collected stated investment choices. As in the first-stage experiment, parents were presented with a series of hypothetical scenarios. But in this stage, instead of asking respondents to report what they believed the outcome would be, they were asked to select their favourite allocation choice. This approach, which relates to contingent valuation methods used in the field of marketing research, has been recently used in economics to study preferences for workplace attributes, university choices, marriage markets, saving behaviour and labour force participation (Mas & Pallais (2017); Wiswall & Zafar (2018); Delavande & Zafar (2019); Adams-Prassl & Andrew (2020); Ameriks, Briggs, Caplin, Shapiro, & Tonetti (2020); Ameriks, Briggs, Caplin, Lee, et al. (2020)).

In the experiment, respondents were presented with a representative family deciding how to distribute educational resources between their children. The resources being allocated were described in terms of monetary investment made by the family in each child in terms of school fees, private tuition, stationary, books and other school related expenditures (these are all expenditure items that are familiar to the respondents in the sample as they are part of their educational budget). Similarly to the first-stage experiment, the survey script emphasised that these were *long-run* investments that would help the children acquire new skills and progress through their educational careers (the exact wording of some relevant questions in the survey is presented in Online Appendix C).

While the equilibrium allocation rule in (8) does not allow to separately identify the primitive parameters in the demand function, it does provide predictions about parental behaviour that I use to guide the design of the hypothetical scenarios in the experiment. In particular, according to (8), as the difference in children's baseline abilities increase parents would invest more in the higher-ability child when $\gamma > 0$, and do the opposite when $\gamma < 0$. In the former case, the investment strategy is reinforcing, while in the latter it is compensating. Therefore, between scenarios I introduced exogenous variation in the *difference* in baseline ability between the two children. As in the case of beliefs, while one child was described as being *"among the top three students in his/her class"*, the other child was described as either being *"among the bottom three students in his/her class"* or as *"an average student in his/her class".* After presenting each scenario, respondents were asked to distribute investment inputs between the two children, by physically allocating some tokens to each child in a labelled account.

To alleviate concerns that the experiment made salient to the parents some specific choices, enumerators did not emphasise what different allocations would achieve in terms of total returns or difference in outcomes. Once parents made their choices, surveyors recorded the answer, collected the tokens and moved on to the next scenario. All hypothetical scenarios were presented with the help of visual aids similar to those used to elicit parents' beliefs. To ensure understanding, two practice scenarios in which parents had to allocate tokens according to a well defined allocation were presented at the beginning of the experiment. If parents could not correctly identify the practice allocations, surveyors continued explaining how to do it.[17]

Because previous research has demonstrated that for poor households an important constraint to invest in children's education is the availability of material resources (e.g., Lochner & Monge-Naranjo (2012)), I test whether household resources also matter to explain allocations by exogenously varying the total level of resources across scenarios. Finally, to understand whether *on average* investments differ between sons and daughters, respondents were randomised in two groups, where one group would see in each scenarios two children who were boys, while the other group would see one girl and one boy.

**Identifying the reduced-form demand function.** Comparing allocations between scenarios, I identify the causal relation between child baseline ability and parental investment, by exploiting the experimentally-induced variation in children's endowments. The experimental design has several important advantages over the use of observational data. First, in the experiment I can abstract from unobserved child specific preferences, as well as other features of the environment that are unobserved by the researcher and might drive the relation between child endowment and parents' investment, as these are held fixed by design of the hypothetical scenarios. Second, unlike observational data, the experimental allocations are both *private*, in the sense that they cannot be shared among children, and *assignable*, that is specific to one individual child known to the researcher, so that I do not need to impose any assumption on how investments are shared or consumed by individual children.[18] Third, the identification strategy is robust to parents having inaccurate beliefs about their children's baseline ability (Dizon-Ross (2019)), as these abilities are precisely described to the respondent in each scenario. Finally, by using individual level variation, the method has a considerably higher statistical power to detect differences in educational investments than the standard approach of using exogenous changes in child characteristics across respondents.

Formally, to test whether parents' investment reinforce or compensate baseline differences between

[17] Virtually all parents could correctly allocate tokens in the practice allocations.

[18] This is also important because standard household surveys typically collect information on parents' educational investment at the level of the household as a whole (e.g., educational expenditure on children), rather than at the level of the individual child.

children, I estimate the following empirical specification by OLS:

$$s_{i,j} = \beta_0 + \beta_1 diff_j + \eta_i + u_{i,j} \tag{10}$$

where $i$ indicates the respondent and $j$ the scenario, $s_{i,j}$ is the share of total resources allocated to the higher-ability child in scenario $j$, and $diff_j$ is a dummy variable that equals one in the scenarios where the difference between the two children's baseline ability is high. The sign of $\beta_1$ pins down whether parents' investment is reinforcing ($\beta_1 > 0$) or compensating ($\beta_1 < 0$). To test whether household resources matter to explain allocations, I expand equation (10) and estimate:

$$s_{i,j} = \beta_0 + \beta_1 diff_j + \beta_2 res_j + \beta_3 diff_j \times res_j + \eta_i + u_{i,j} \tag{11}$$

where $res_j$ is a dummy variable that takes value one if in scenario $j$ resources are high. The sign of $\beta_3$ identifies if reinforcement (compensation) is weaker when resources are higher (lower) ($\beta_3 < 0$ or $\beta_3 > 0$). Finally, by comparing investment allocations between the group of respondents that saw two boys in the hypothetical scenario, with the group of respondents that saw one boy and one girl (holding fixed other child characteristics), I study the role of child gender.

### 4.3 Combining Measures to Identify Preferences

While estimates of equations (10) and (11) identify whether parents' investment reinforce or compensate differences in children's baseline abilities, without further assumptions on parental beliefs one can not back-out parental preferences for intra-household inequality. This is easily illustrated by looking again at equation (8), which is reported here for convenience and abstracts from child specific preferences that are held fixed by design in the hypothetical scenarios:

$$\log\left(\frac{I_i^*}{I_j^*}\right) = \gamma \log\left(\frac{A_i}{A_j}\right) = \frac{a\rho}{1-b\rho}\log\left(\frac{A_i}{A_j}\right) \tag{12}$$

By combining experimental data on beliefs and choices, I identify parental preferences for intra-household inequality. The intuition for the identification result is simple. A regression of expected child outcomes on investment and ability identifies the parameters of the perceived production function ($a$ and $b$). The reduced-form parameter $\gamma$ is identified from data on experimental allocations. Once these parameters are identified, parents' preferences can be recovered as:

$$\rho = \frac{1}{a} \times \left[\frac{1}{\gamma} + \frac{b}{a}\right]^{-1} \tag{13}$$

A consistent estimator for $\rho$ can be obtained by replacing the parameters in (13) with the corresponding OLS estimates from equations (9) and (10).[19,20] For this identification strategy to work it is important

[19] Consistency of the estimator for $\rho$ follows from the consistency of the OLS estimator for the parameters in (9) and (10), using the continuous mapping and the Slutsky theorems. I obtain standard errors and confidence intervals for the preference parameter using bootstrap methods.

[20] See Appendix B.3 for the derivation of equation (13).

that parents act upon their beliefs when making their investment decisions. Recent experimental evidence points to the fact that this is indeed the case, as changing parents' beliefs has been show to causally impact the type and level of investments that parents select for their children (Dizon-Ross (2019); List, Pernaudet, & Suskind (2021)).

## 5 Data and Descriptive Statistics

The experiment was conducted with 504 families with children living in the urban slums of Cuttack, Odisha, India. The state of Odisha is located in Eastern India and is one of the poorest, with 33 percent of its residents living below the poverty line (Reserve Bank of India (2017)).

The data collection was part of a long-run follow-up of a cluster randomised controlled trial of a psychosocial stimulation intervention targeted at disadvantage children. In 2013, a sample of young and poor women with children (aged 10 to 20 months then) was identified through a door-to-door census. Of these 46 percent lived below the poverty line.[21] Households were then randomised in a treatment and a control groups. The treatment group participated in home visiting activities aimed at improving mother-child interactions and promote child development.[22] Appendix Tables A.1 and A.2 show that there are no effects of the treatment allocation on parents' beliefs or preferences, and that the main results are robust to the exclusion of households in the treatment group. The results from the first follow-up also showed that there were no improvements in maternal knowledge of child development in this sample (Andrew et al. (2019)). Similarly, Attanasio, Cunha, & Jervis (2019) show that the same intervention did not change parents' beliefs in the long run in a sample of poor parents in Colombia. For the main results, I thus ignore the treatment allocation and pool the treatment and control groups together, but show that the results are robust when only considering households in the control group.

In 2019, we aimed at re-interviewing all households in the original sample to study whether there were sustained benefits from the intervention (these results are not reported in this paper). To increase the sample size (only for the purpose of the present study), in larger slums one or two neighbours of randomly selected households from the original experimental sample were also interviewed. To take part to this study, the neighbour household had to have at least one child of the same age as "target" children from the original study (i.e., between 6 and 8 years old at the time of this study).

Survey respondents were for the most part children's female primary caregivers, who were usually their mothers. The survey experiment took place in respondents' homes, during the caregivers' endline survey and, whenever possible, in a quiet and private environment. The endline survey also collected household characteristics, and, separately for each child, detailed information on their education, parents' human capital investments (including child-specific educational expenditure, time investments by the parents such as time spent helping children with homework), and children's time use.

Table 1 reports the summary statistics for key variables in the sample. It shows that this is an economically and socially disadvantaged sample: over 50 percent of children's primary caregivers did not complete lower primary education, and just over 30 percent of households are attached to the sewage

[21] The Rangarajan committee urban poverty line is INR 47 per household member per day.

[22] Andrew et al. (2019) report details of the intervention, evaluation design and short-run results. The authors find a positive effect on the cognitive development of target children at the end of the intervention.

| | Mean | S.D. |
|---|---|---|
| *A. Household characteristics* | | |
| Primary caregiver did not complete primary | 0.508 | 0.500 |
| Primary caregiver age | 27.933 | 6.216 |
| Household size | 6.512 | 3.285 |
| Number of children | 2.296 | 0.930 |
| Household owns dwelling | 0.712 | 0.453 |
| Number of rooms | 2.766 | 2.278 |
| Household is attached to sewage system | 0.312 | 0.464 |
| Yearly food expenditure (thousands)† | 71.463 | 49.788 |
| *B. Children's characteristics* | | |
| Child age | 7.438 | 3.510 |
| Child is male | 0.482 | 0.500 |
| Yearly educational expenditure per child (thousands)† | 6.662 | 9.555 |
| *C. Household members' characteristics (excluding children)* | | |
| Household member age | 34.195 | 15.879 |
| Household member is male | 0.478 | 0.499 |
| Total number of households | 504 | |
| Total number of children | 1196 | |
| Total number of individuals | 3282 | |

*Notes:* This table presents the summary statistics for the sample. Panel A reports primary caregiver's and household's characteristics, Panel B the characteristics for children and Panel C the statistics for all household members (excluding children). † Indicates expenditure in thousands of INR. Educational expenditures includes school fees, uniforms, textbooks, stationary and after-school tutoring. The exchange rate was 71.43 INR : 1 USD at the time of the study.

Table 1: Summary Statistics

system. While most households own the dwelling they live in (71 percent), these are usually small in size, with on average two rooms for more than six household members. Families in the sample are relatively young as shown by the average age of the respondent of 28 years old. There are on average two children in each family, and their average age is 7.5. Therefore, for most parents distributing resources between *two* children as in the hypothetical scenarios is relevant and realistic as this is the *actual* choice they face. Table 1 also shows that among children the percentage of boys is 48 percent, which implies a balanced sex ratio (this is also true if we consider all household members and not children specifically). As a reference, in 2019 the national sex ratio was 940 girls per 1000 boys in India, and 978 girls per 1000 boys in Odisha (Indian Census (2011)).

## 6 Results

This section discusses the results and is organized as follows. Section 6.1 presents the results on parental subjective beliefs about the human capital production function. The experimental results on parents'

| | Perceived earnings at age 30 (log) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| High ability | 0.831*** | 0.911*** | 0.848*** | 0.768*** | 0.404*** |
| | (0.021) | (0.031) | (0.032) | (0.024) | (0.040) |
| High Investment | 0.252*** | 0.252*** | 0.189*** | 0.189*** | 0.146*** |
| | (0.011) | (0.011) | (0.014) | (0.015) | (0.015) |
| Boy | | 0.160*** | 0.160*** | | |
| | | (0.044) | (0.044) | | |
| Ability × Investment | | | 0.126*** | 0.126*** | 0.128*** |
| | | | (0.018) | (0.019) | (0.019) |
| Belief about child education | | | | | 0.159*** |
| | | | | | (0.017) |
| Mean outcome | 30120 | 30120 | 30120 | 30120 | 30120 |
| $R^2$ | 0.361 | 0.367 | 0.369 | 0.774 | 0.800 |
| Observations | 3960 | 3960 | 3960 | 3960 | 3960 |
| Respondent fixed effects | | | | ✓ | ✓ |

*Notes:* The outcome variable is log-earnings of the child at age 30 as perceived by the respondent. Columns 1 to 3 display the OLS results. Columns 4 to 5 further include family fixed effects. Robust standard errors clustered at the respondent level are reported in brackets. *High ability* is a dummy variable that takes value 1 if in scenario $j$ the child has an high academic ability, *High investment* is a dummy variable that takes value one if in scenario $j$ the level of investments is high, and *Boy* is a dummy variable equal to one if the child is a boy. *Belief about child education* is the educational attainment respondents believe the child would achieve in scenario $j$. $^*$ denotes 10% significance, $^{**}$ denotes 5% significance, $^{***}$ denotes 1% significance.

Table 2: Perceived Production Function

allocations and preferences are presented in section 6.2. Section 6.3 relates preferences elicited in the experiment to actual educational investments made by parents in their own children.

### 6.1 Beliefs

I present the estimates of equation (9) in Table 2, where the outcome variable is (log) child earnings at age 30 as expected by the parents (similar results for educational attainment are reported in Appendix Table A.3). I start by regressing perceived earnings on a dummy for high baseline ability and a dummy for high investment in column 1. I subsequently control for child gender and for the interaction between ability and investment (columns 2 and 3). In column 4, I also include respondent fixed effects. Finally, in column 5, I control for child educational attainment (as expected and reported by the respondent).

Parents perceive the returns to baseline ability to be large, with an expected increase in earnings of 76-91 percent (columns 1 to 4). At the sample mean of expected earnings (30,120 INR) this corresponds to an increase of roughly 24,000 INR. I discuss the magnitude of this coefficient below. The coefficient on child baseline ability decreases by almost 50 percent when I control for expected years of schooling, as reported by the respondents (column 5). This is because, as show in Appendix Table A.3, parents believe that higher-ability children would achieve on average two more years of schooling compared to lower-ability children. In turn, one year of schooling is associated with an increase in expected earnings of 15.9 percent (column 5 of Table 2). This is not a causal effect (as education was not randomised across scenarios) but suggests that schooling is one likely mediator for the effect of child ability on earnings.

Turning to the perceived returns to investment, column 2 shows that parents believe that increasing educational expenditure from the 10th to the 90th percentile in the sample, would increase child earnings at age 30 by 25.2 percent. This coefficient slightly decreases when controlling for the interaction between ability and investment (columns 3 and 4), and further declines when controlling for expected years of schooling as reported by the respondents (column 5).

Finally, the results in column 3 imply that baseline ability and investment are perceived as complements: parents believe the returns to investments to be 12.6 percent higher for the higher-ability child compared to the low-ability child. This perceived complementarity generates an incentive for parents to reinforce initial differences between children if they seek to maximize the returns from their investment. These results are similar to Boneva & Rauh (2018), who report that parents in the UK perceive positive returns from investments and ability and that the two inputs are complements.[23]

**Benchmarking perceived returns.** Table 2 reports the coefficients associated with a binary increase in the relevant input (i.e., a change from a *low* level of the input to a *high* level). As such, they can not be easily interpreted or compared. To ease interpretation and comparability, I convert these coefficients in terms of a one-standard-deviation increase in the relevant input. This exercise reveals that parents perceive a one-standard-deviation increase in ability to increase earnings by 15 percent. Similarly, a one-standard-deviation increase in investments is expected to boost earnings by 28 percent.

To put these figures into perspective, I contrast them with expected gender-gap in earnings. In the sample, parents expect boys to earn on average 16 percent more than girls at age 30 (columns 2 and 3 of Table 2). Interestingly, this figure is not far from the actual gender-gap in urban workers' earnings of 22 percent (ILO (2018)). I terms of beliefs related to child gender, I also find that while parents believe that girls on average will command less resources than boys as adults (as implied by the coefficient on child gender in Table 2, and show in Appendix Table A.4), they do not perceive the returns to ability or investment to substantially differ between girls and boys (see Appendix Tables A.4, A.5 and A.6). These findings imply that respondents do not perceive the technology of skills formation to differ by gender, but are suggestive of the fact that parents incorporate in their beliefs the social norms prevailing in their community, and reflecting the differential opportunities that men and women face in the labour market. These results are also similar to Boneva & Rauh (2018) who find significant differences in average perceived earnings but not in the perceived returns of different inputs by gender in the UK.[24]

**Beliefs heterogeneity.** The estimates in Table 2 represent *average* parental beliefs. To uncover heterogeneity between respondents (that I later use in the analysis), I follow Boneva & Rauh (2018)

[23] Results for educational attainment (Appendix Table A.3) follow a quantitative similar pattern, except that the interaction between child baseline ability and parents' investments is not statistically different from zero for this outcome.

[24] Using the estimates from Attanasio, Meghir, & Nix (2020) it is also possible to compare parents' perceptions with estimates of the actual human capital production function in India. Attanasio, Meghir, & Nix (2020) finds that a one-standard-deviation increase in child baseline cognitive skills increase next period cognitive skills by 0.6 of a standard deviation, and a one-standard-deviation increase in investment increases child human capital by 0.2 standard deviation at age 8. They also find that investment and baseline development are complements.

and construct an individual-specific measure of perceived returns. For example, I compute individual perceived returns to investment as the difference between respondents' expected earnings reported in the scenarios in which investment is high and the corresponding scenarios in which it is low (i.e., holding fixed other characteristics of the child and the scenario), and average these differences across scenarios. I plot the empirical cumulative distribution of perceived returns in Appendix Figure A.3. Panel A displays the distributions of perceived returns to child baseline ability. The figure reveals a substantial variation in perceived returns across respondents: the 10th percentile is 0.33 and the 90th percentile is 1.19. By comparing expected earnings in high and low investment scenarios, while holding child ability fixed, I also compute individual perceived returns to investment. The distribution of these perceived returns is shown in panel B of Appendix Figure A.3, and also shows substantial heterogeneity: the 10th percentile is 0 and the 90th percentile is 0.48. Figure A.3 further shows that, consistently with the findings form Table 2, the distribution of perceived returns to investment for the the higher-ability child first order stochastically dominates that of the lower-ability child.[25]

### 6.2 Investment Choices and Preferences

Table 3 reports the estimates of equations (10) and (11). I start by running the model without respondent fixed effects (columns 1 and 3) and then add them in (columns 2 and 4). The coefficient in column 1 shows that, as the difference between children's baseline ability increases, parents re-allocate and devote a significantly larger share of resources to the higher-ability child. The point estimate implies a 7.8 percentage points increase in the share of resources allocated to the higher-ability child, which at the sample mean corresponds to a 14 percent increase in the share of resources devoted to this child (equivalently, a one-standard-deviation increase in the difference between children's baseline abilities leads to a 9.7 percent increase in the share of resources allocated to *Child H*). This result implies that parents' investment is *reinforcing*.

Table 3 also shows that household resources play a role in explaining parents' allocations. The results in column 3 show that reinforcement is stronger when resources are lower. This is captured by the negative and statistically significant coefficient on the interaction between children's ability and resources. Specifically, when resources are low the share allocated to the higher-ability child is 10.2 percentage points higher in scenarios where the ability difference is large compare to when it is small. The gap between children is halved when resources are high. This result highlights the role that household constraints have to explain the allocation of human capital investments between children. The findings are consistent with Behrman (1988), who shows that parents favour better endowed children in the lean season, and more generally with the idea that "*discrimination is stronger in a time of crisis*" (Duflo (2005)). Therefore, it seems plausible that relaxing resource constraints could contribute to close investments gaps between children, potentially also resulting in lower intra-household inequality in outcomes.

Finally, Table 3 shows no evidence that investment choices depend on the gender of the child.[26]

[25] I correlate individual perceived returns with observable characteristics in Figure A.5 and find that the education level of the primary caregiver predicts higher perceived returns to investment. Appendix Tables A.1 show that there is no significant effect of the psychosocial stimulation intervention on parents' beliefs.

[26] I also tried estimating equations (10) and (11) separately for the two different groups defined based on the gender of

| | Share of resources to child $H$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Difference in ability | 0.078*** | 0.078*** | 0.102*** | 0.102*** |
| | (0.005) | (0.006) | (0.008) | (0.010) |
| High resource | | | 0.028*** | 0.028*** |
| | | | (0.007) | (0.008) |
| Difference in ability × High resources | | | -0.048*** | -0.048*** |
| | | | (0.008) | (0.010) |
| Boy | -0.001 | | -0.001 | |
| | (0.008) | | (0.008) | |
| Mean outcome | 0.541 | 0.541 | 0.541 | 0.541 |
| $R^2$ | 0.078 | 0.535 | 0.085 | 0.542 |
| Observations | 1980 | 1980 | 1980 | 1980 |
| Respondent fixed effects | | ✓ | | ✓ |

*Notes:* The outcome variable is the share of total resources invested in child $H$. This variable ranges from 0 to 1. Columns 1 and 3 display the OLS results, while columns 3 and 4 further includes family fixed effects. Robust standard errors clustered at the respondent level are reported in brackets. *Difference in ability* is a dummy variable that takes value 1 if in scenario $j$ the difference between the two children's academic ability is large and zero otherwise, *High resources* is a dummy variable that takes value one if in scenario $j$ the level of resources is large and zero otherwise, and *Boy* is a dummy variable that takes value one if the respondent was randomised in seeing two boys and zero if the respondent was randomised in seeing one boy and one girl. $^{*}$ denotes 10% significance, $^{**}$ denotes 5% significance, $^{***}$ denotes 1% significance.

Table 3: Intra-household Allocation of Resources

Although the previous literature does not always find evidence of differential treatment of boys and girls, recent work on India shows that boys are breastfed longer (Jayachandran & Kuziemko (2011)), and receive more childcare time early in life (Barcellos, Carvalho, & Lleras-Muney (2014)). To interpret the results in Table 3, one has to keep in mind that the input being allocated in the experiment is educational *expenditure*. Consistently with my findings, previous research has found no evidence of parents spending differently on boys and girls (Deaton (1989, 1997)). Similarly, there is little evidence that girls of primary school age receiving less human capital investments compared to boys in urban Odisha, suggesting that son preferences might be less prevalent in the context of this study compared to other Indian states.[27]

It might then seem odd that parents equally allocate educational investments between sons and daughters, despite them perceiving girls to be able to command less resources as adults (as implied by the results in Table 2). One potential reason that could explain this result is that, when deciding on their daughter's schooling, parents also consider the marriage market returns to girls' education (in addition to the labour market returns). Indeed, recent evidence suggests that a key motivation

the two children, and found very similar results.

[27] In terms of educational investments, in urban Odisha school attendance is the same for boys and girls in the age groups 6-10 years and 15-17 years, and slightly higher for girls than boys in the age group 11-14 years (81% of girls compared with 78% of boys). Similarly, in terms of health investments and outcomes, the infant and under-five mortality rates are 23-26 percent higher for boys than for girls. Among surviving children, girls and boys are about equally likely to be undernourished. Girls are also more likely than boys to be fully vaccinated (55% of girls, compared with 49% of boys) (IIPS (2001, 2008); Padhi (2001)).

for investing in daughters' education is a substantial perceived marriage market return to schooling (Adams-Prassl & Andrew (2020); Ashraf, Bau, Nunn, & Voena (2020)). In any case, as show in Table 2, perceived returns to investment are not heterogeneous by child gender, so the result in Table 3 is not inconsistent with parents simply allocating resources based on their perceived returns.

**Parents' preferences.** As discussed earlier, using the experimental allocations to regress parents' investment on child ability identifies the reduced-form parameter $\gamma$. This comprises both parental preferences for inequality and their perceptions about the production function. Using parents' beliefs from the first stage experiment, I identify preferences using the procedure outlined in Section 4.3.

I find that the value of $\rho$ that reconciles choices with parents' beliefs is positive and statistically significant at the 99 percent confidence level, implying that in this setting parents' investment choices are primarily driven by efficiency considerations rather than by inequality concerns over final outcome: the point estimate is 0.449, with an associated standard error of 0.041. Interestingly, the estimated parameter is very close to the parameter value estimated by Behrman (1988) using observational data from India (0.47). The estimated $\rho$ is also statistically different from 1 at the 99 percent confidence level, suggesting that while parents weight relatively more efficiency than inequality-concerns, they are not pure returns-maximizers when investing in their children (as would implied by a value of $\rho$ equal to 1). This results is consistent with parents also placing some value on equalizing inputs (rather than outcomes), as suggested by Berry, Dizon-Ross, & Jagnani (2020), although this preference can not be directly identified in this experiment.

As for beliefs, I also study heterogeneity in preferences. I plot the empirical cumulative distribution function of individual preferences in Appendix Figure A.4 and find that $\rho$ is positive for all families in the sample. However, there is heterogeneity across respondents, so that some families have an higher value of $\rho$ than others (i.e., they are less inequality averse). I use this heterogeneity to classify families as *low* and *high $\rho$ types* by splitting the sample at the median value of the empirical distribution of $\rho$, and investigate how elicited preferences related to non-experimental choices in the next section.[28,29]

## 6.3 Stated and Revealed Preferences

I next investigate the relevance of the results outside the experiment, by considering whether elicited preferences reflect *real world* behaviour. To answer this question, I exploit a key feature of the survey, in which parents were asked detailed information on educational investments separately for each of their children, both in terms of monetary investment (e.g., educational expenditure) and time investment (e.g., help with homework). This represents an important improvement over standard household surveys, which typically collect investment information at the level of the household as a whole (e.g., educational expenditure for all children), or only on child-specific health inputs (e.g., breastfeeding or vaccinations which are common in large scale household surveys), and allows me to study how parents

[28] To construct this figure, I use individual perceived returns from the previous section rather than average beliefs.

[29] I correlate parental preferences with observable characteristics in Figure A.6 and find that households attached to the sewage system are more inequality averse, while families with more children are less inequality averse. This latter result is consistent with the model extension considered in Appendix B.1, which predicts that fertility increases with $\rho$. Appendix Table A.2 shows that there is not significant effect of psychosocial stimulation intervention on parents' allocations.

allocate key human capital inputs such as educational resources and time.

I use this rich information on current investments and relate it to the ability of respondents' own children. To measure child ability, I rely on the following survey question: "*Using the scale, can you please show me how intelligent do you think "child" is? In general, not only in school. If you think that "child"'s intelligence is extremely good you should score 10, while if you think that "child"'s intelligence is very poor you should score 0.*". This questions captures a belief held by parents about their children's ability, which might or might not be accurate (Dizon-Ross (2019)). Importantly, what matters to understand intra-household allocations is whether these beliefs (more precisely the difference in beliefs between two children) explain parental investment, and not whether these beliefs are correct.

The results are presented in Table 4. I start by regressing the difference in educational expenditure between two children on the difference in their ability, controlling for other child characteristics. The results in column 1 suggest a positive and significant relation between child ability and investment. In particular, the point estimate implies that a 10 percent increase in the difference between children's abilities is associated with an increase in the educational expenditure gap of 290 INR in favour of the higher-ability child. At the sample mean, this corresponds to 3.8 percent of total yearly educational expenditure. Appendix Table A.7 breaks this down by expenditure categories: total school fees, uniforms, textbooks, stationary and after-school tutoring. The point estimate are positive across all the outcomes considered, and the largest effect is on after-school private tutoring, for which parents spend and additional 234 INR on their higher-ability child.

Next, I turn to the question of whether elicited preferences are predictive of actual choices. To answer this question, I exploit the heterogeneity in preferences reported in the previous section and classify families as more or less inequality averse (depending on whether the respondents' estimated $\rho$ is above or below the sample median, where a value above the median mean lower inequality aversion). Column 2 and 3 report the results. I find that respondents identified as less inequality averse in the experiment, systematically make more unequal allocations when it comes to distribute actual resources. In particular, the point estimate in column 3 is over twice as large as that in column 2 and statistically different from zero. The point estimate implies that a 10 percent increase in the difference between children's ability increases the educational expenditure gap between the higher- and lower-ability child by 509 INR. This corresponds to 6.8 percent of the yearly educational expenditure. On the other hand, for families classified as more inequality averse this figure is 240 IRN (not statistically significant), which corresponds to 3 percent of total yearly educational expenditure.

The fact that my experimentally elicited measure of parental preferences maps into actual investment behaviour is reassuring, and adds credibility to the research design and to the use of strategically-designed survey measures based on hypothetical scenarios to identify primitive parameters of interest. This also aligns with recent evidence pointing to the fact that the two approaches of using stated or actual choices yield similar conclusions in a variety of contexts, especially when the hypothetical scenarios are realistic and relevant for the respondent (Mas & Pallais (2017); Wiswall & Zafar (2018)), and with the results in Berry, Dizon-Ross, & Jagnani (2020) who find that parents make similar investment decision in a incentivized real-stakes setting and in a hypothetical online survey experiment.

The remainder of Table 4 report the results for additional investment measures. Columns 4 to 9

| | Educational expenditure† | | | Child work | | | Private school | | | Time index | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All | $\rho_L$ | $\rho_H$ | All | $\rho_L$ | $\rho_H$ | All | $\rho_L$ | $\rho_H$ | All | $\rho_L$ | $\rho_H$ |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Ability | 290.594* | 240.393 | 509.987** | -0.052*** | -0.035*** | -0.079*** | 0.007 | 0.004 | 0.011 | 0.212*** | 0.207*** | 0.223*** |
| | (171.511) | (241.313) | (255.433) | (0.009) | (0.011) | (0.014) | (0.006) | (0.008) | (0.010) | (0.019) | (0.024) | (0.032) |
| Child controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean | | 7437 | | | 0.15 | | | 0.11 | | | 0 | |
| Observations | 582 | 281 | 301 | 504 | 247 | 257 | 548 | 264 | 284 | 417 | 201 | 216 |

*Notes:* The outcome is the difference in investment between two children, as measured by the outcome in the column header. Columns 1, 4, 7 and 10 report the results in the full sample, while the remaining columns report separate results for two separate sub-samples as defined by their inequality aversion ($\rho_L$ means higher inequality aversion). These two groups are defined based on whether the estimated $\rho$ falls above or below the sample median. Child controls include gender and age. Robust standard errors clustered at the family level are reported in brackets. † in INR. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table 4: Actual Parental Investments

show the results for whether the child works (this includes both remunerated and non-remunerated work activities performed by the child) and the type of school attended (private vs. public). I find that higher-ability children are more likely to attend a private school (although the point estimate is not statistically significant). At the same time they are 5.2 percentage points less likely to work, which amounts to a 35 percent decrease in the likelihood of working. Moreover, similarly to the results for educational expenditure these differences in investments across higher- and lower-ability children are more pronounced in less inequality averse families than in more inequality averse ones.[30]

Finally, columns 10 to 12 of Table 4 report the results for a time investment index, constructed using items from the HOME inventory (Bradley, Caldwell, Rock, Hamrick, & Harris (1988)). Overall, I find that higher-ability children score 0.21 standard deviations higher on the aggregate time investment index, but there is no large difference between more or less inequality averse parents in terms of this outcome. Appendix Table A.8 report results on different sub-components of the time index, showing that parents spend more time engaging with their higher-ability child on homework and other school-related activities. For example, parents are 2.5 percentage points more likely to help their child with school work (column 5). The results also suggest that parents do not spend more time with their higher ability child across the board (e.g., they spend the same amount of time playing with their children as shown in column 4), but specifically tailor school-related activities to their children's ability. These additional results suggest that parents' investment decisions might extend beyond educational expenditure, with potentially important long-term effects for children's wellbeing.

# 7 Policy Experiment and Welfare

Several studies show that information provision affects individual decision-making across different domains (Jensen (2010); Dupas (2011), Liebman & Luttmer (2015); Fitzsimons, Malde, Mesnard, & Vera-Hernández (2016); Dizon-Ross (2019); Hjort, Moreira, Rao, & Santini (2021)). In this section, I explore the implications of my results to understand the effects of a policy designed to improve child human capital. Specifically, I consider the effects of an intervention that affects parents' beliefs about the returns to investments in children, while holding fixed the educational budget. Similar interventions have become increasingly popular in recent years both in developed and in developing countries, and are often targeted at disadvantaged families under the presumption that one of the reasons they under-invest in their children is the low perceived benefits of parental investment (e.g., York, Loeb, & Doss (2019); List, Pernaudet, & Suskind (2021); Duncan, Kalil, Mogstad, & Rege (2022)). Although these programmes are usually delivered at the household level, they typically have one target child. Consequently, evaluations of such interventions typically collect data on investment and outcomes for the target child only. Here, I consider the effects of the policy on parents' investment in *each* child, on children's outcomes, as well as on household and individual child welfare. I assume that child welfare corresponds to her final level of human capital, while parents' welfare corresponds to their utility.

For the counterfactual exercise, I assume that the family has two children, and there are no child-

[30] Appendix Table A.9 shows that the main results are qualitatively very similar if household in the treatment group are dropped from the estimation sample. This is consistent with the fact that there are no effects of treatment allocation on parents' beliefs or preferences (Appendix Tables A.1 and A.2).

| | $U$ (1) | $I_H$ (2) | $I_L$ (3) | $H_H$ (4) | $H_L$ (5) |
|---|---|---|---|---|---|
| *Panel A:* $\rho = 0.449$ | | | | | |
| Before | 12.151 | 6.469 | 3.531 | 19.449 | 3.113 |
| After | 12.193 | 7.251 | 2.749 | 21.556 | 2.484 |
| *Panel B:* $\rho = 0.384$ | | | | | |
| Before | 12.156 | 6.258 | 3.742 | 18.878 | 3.280 |
| After | 12.174 | 6.804 | 3.196 | 20.354 | 2.846 |
| *Panel C:* $\rho = 0.521$ | | | | | |
| Before | 12.518 | 6.699 | 3.301 | 20.071 | 2.930 |
| After | 12.613 | 7.796 | 2.204 | 23.006 | 2.037 |

*Notes:* This table reports the effects of a policy experiment that corrects parents' beliefs about the productivity of investment. The first column reports total household utility, the second and third columns report parents' investment in *Child H* and *Child L*, and the last two columns report children's welfare, which corresponds to their level level of human capital separately for the higher- and lower-ability child. Panel A reports the results when $\rho$ is set to its estimated value of 0.449 from section 6.2, while in Panels B and C report the results setting $\rho$ to the lower and upper bound of the 95% confidence interval. For each panel, the first and second rows compare before and after the policy is implemented.

Table 5: Policy experiment

specific preferences, so that the utility function is given by:

$$U(H_1, H_2) = (H_1^{\rho} + H_2^{\rho})^{\frac{1}{\rho}}$$

These assumptions are made only for simplicity, and the main results in this section do not depend on these assumptions. The two children have different baseline abilities, so that $\theta_1 = \theta_H$ and $\theta_2 = \theta_L$, with $\theta_H > \theta_L$. The policy increases parents' *perceived* returns of investment $a$, aligning it to its *true* productivity $\alpha$, which is assumed to be larger so that $a < \alpha$ (this assumption in consistent with the evidence in Cunha, Elo, & Culhane (2013) and Boneva & Rauh (2016)). In the simulations, I keep fixed all other model parameters; in particular, I set the parents' preference parameter $\rho$ equal to its estimated value from Section 6.2: $\rho = 0.449$. The results are shown in Panel A of Table 5.

Table 5 shows that, as the policy corrects parents' distorted beliefs about the child human capital production function and parents' re-optimize accordingly, total household welfare $U$ goes up after the policy is implemented (column 1). The magnitude of the effect of the policy on parents' welfare is comparable to a 1% increase in the total educational budget $y$.[31] At the same time, because the intervention increases the perceived returns to investment, parents adjust their investment behaviour: Compared to the pre-policy period, investment in the higher-achieving child increases and investment in the lower-achieving child decreases ($I_H$ and $I_L$ in columns 2 and 3). As a result, the policy improves the welfare of the higher-ability child, but decreases that of the lower-ability child ($H_H$ and $H_L$ in columns 4 and 5). These results are robust to using different values of parents' inequality aversion, as shown in Panels B and C of Table 5, where I report the same set of results, setting $\rho$ equal to the lower and upper bounds of the 95% bootstrap confidence interval from Section 6.2.[32]

[31] This is computed using estimates from the model and solving for the level of income that generates the same increase in household utility.

[32] Re-optimization from the parents following the policy will occur for all values of $\rho$ that are different from zero.

Taken together, the results of the counterfactual policy experiment highlight the importance of taking parents' endogenous responses into account when considering the effects of policies designed to improve children's welfare. To the extent that parents are the ultimate decision makers, it is necessary to understand how they allocate resources to individual children and what determines their behaviour in order to predict the effects of policies and understand their impacts on individual children's wellbeing. These findings suggest that, while some policies might be welfare improving for the household as a whole, they might conceal important distributional impacts, so that some children might benefit while other might be worse off. These considerations are often overlooked when thinking about the design and evaluations of programmes targeting the home environment, but might be important to gain a better understanding of their (lack of) impacts.[33]

Finally, it is important to highlight that in the policy experiment the educational budget is held fixed. However, it could be that providing information also leads parents to increases total educational expenditure. In that case, while the gap between children's investment would increase the level of investment in each child could also increase. Therefore, the impacts of the policy will ultimately depend on the household's ability to adjust its educational expenditure, suggesting that the intervention could have different effects for poorer (resource-constrained) families and wealthier ones.

# 8 Conclusions

This paper studies the role of parents' human capital investments as a determinant of intra-household inequality in child outcomes. I document that, across a large set of developing countries, within household variation explains as much as 50 percent of overall inequality in children's educational attainment. By looking at the human capital distribution within the family, I then show that while the human capital of high achieving children stays constant as family size increases, the human capital of children at the bottom of the achievement distribution steeply declines with family size. I argue that these patterns are consistent with a *behavioural* origin underlying intra-household inequality in child outcomes, and specifically with the differential treatment of children in terms of human capital investments.

To study the role played by parents' educational investments to explain this inequality, I design and implement a survey experiment, motivated by a simple model household behaviour. I develop new theory-driven survey measures based on hypothetical scenarios that allow me to separately identify parental beliefs about the human capital production function *and* their preferences for inequality in children's outcome, as well as study the role of household resources. I then complement these strategically-designed instruments with available behavioural data to validate my experimental strategy. I implement the experiment with a sample of 500 poor households with children in urban Odisha, India.

Several key results emerge from this study. First, I find that parents perceive child baseline ability and investments to be highly productive, and to be complements in the production of human capital, so that parents should invest more in higher-ability children if they want to maximize the returns from

[33] For example, York, Loeb, & Doss (2019) find that a text messaging intervention for parents of preschool children had positive effects on parents' investment. However the same intervention showed no significant benefits in kindergarten (Doss, Fahle, Loeb, & York (2019)).

their investments. Second, I find that parents have a low aversion to inequality over their children's outcomes, and they act upon their beliefs by reinforcing initial differences between children. This suggest that, in this setting, investment choices are driven by efficiency considerations. Third, I show that household resources are important in explaining educational investments, as parents reinforce more strongly when resources are lower. Finally, I demonstrate that experimentally elicited preferences relate to *actual* household behaviour, and that respondents who are identified as less inequality averse in the experiment, systematically invest more unequally in their children, favouring their higher-achieving child.

Taken together, the results indicate that in this setting families act as a reinforcing agent, magnifying ability-based educational inequalities between children. It remains an open question if the results would be similar in different settings or for different types of investments. First, education in India, as in many other developing countries, is better tailored as serving the needs of higher-achieving students (Glewwe & Muralidharan (2016)). In such contexts, it could make sense for parents to match this investment at the household level by investing more in their better-performing child. However, results could be different in educational systems that provide the same educational opportunities to all. Second, parents could make reinforcing and compensating investment across different dimensions, for example by increasing health investment in lower-endowment children (Yi, Heckman, Zhang, & Conti (2015)).

The findings have important implications for policy. First, the fact that parents respond to early levels of child human capital suggest that early interventions can generate both large *direct* positive effects (Heckman (2006)) and have the potential to produce *indirect* effects through parental endogenous investment responses. However, as the the policy simulation shows these endogenous responses might have distributional impacts. To the extent that families are the ultimate decision makers, it is necessary to consider these behavioural responses to understand the impacts of policies on the wellbeing of individual household members. Second, by highlight a link between constraints and intra-household allocations, the findings point to the role that household resources have to explain human capital outcomes within the family. They suggest that that reducing poverty could disproportionately benefit weaker children. Future work should investigate whether relaxing resources constraints is sufficient to improve the human capital of *all* children. In this respect, Barrera-Osorio, Bertrand, Linden, & Perez-Calle (2011) report that parents adjust their investments in response to a conditional cash transfer programme in Colombia by diverting educational resources away from *non-target* children towards *target* siblings. This results is consistent with the fact that the intervention might have made more salient to the parents the returns to invest in some specific children in the family, leaving their siblings at a considerable risk. Studying the intra-household distributional impacts of interventions, and understanding how to incorporate parents' endogenous responses in the design of effective policies should be an important area for future research.

# References

Adams-Prassl, A., & Andrew, A. (2020). *Preferences and beliefs in the marriage market for young brides* (Tech. Rep.). CEPR Discussion Paper No. DP13567.

Adhvaryu, A., & Nyshadham, A. (2016). Endowments at birth and parents' investments in children. *The Economic Journal*, *126*(593), 781–820.

Aizer, A., & Cunha, F. (2012). *The production of human capital: Endowments, investments and fertility* (Tech. Rep.). National Bureau of Economic Research.

Almond, D., & Mazumder, B. (2013). Fetal origins and parental responses. *Annu. Rev. Econ.*, *5*(1), 37–56.

Ameriks, J., Briggs, J., Caplin, A., Lee, M., Shapiro, M. D., & Tonetti, C. (2020). Older americans would work longer if jobs were flexible. *American Economic Journal: Macroeconomics*, *12*(1), 174–209.

Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D., & Tonetti, C. (2020). Long-term-care utility and late-in-life saving. *Journal of Political Economy*, *128*(6), 2375–2451.

Andrew, A., Attanasio, O., Augsburg, B., Day, M., Grantham-McGregor, S., Meghir, C., ... Rubio-Codina, M. (2019). Effects of a scalable home-visiting intervention on child development in slums of urban india: evidence from a randomised controlled trial. *Journal of child psychology and psychiatry*.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*(4), 773–824.

Arcidiacono, P., Hotz, V. J., & Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, *166*(1), 3–16.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2020). Bride price and female education. *Journal of Political Economy*, *128*(2), 591–641.

Attanasio, Cunha, F., & Jervis, P. (2019). *Subjective parental beliefs. their measurement and role* (Tech. Rep.). National Bureau of Economic Research.

Attanasio, & Kaufmann, K. (2009). *Educational choices, subjective expectations, and credit constraints* (Tech. Rep.). National Bureau of Economic Research.

Attanasio, O. (2021). Economics and measurement: New measures to model decision making. *In Econometric Society Presidential Address.*.

Attanasio, O., Bernal, R., Giannola, M., & Nores, M. (2020, September). *Child development in the early years: Parental investment and the changing dynamics of different dimensions* (Working Paper No. 27812). National Bureau of Economic Research. doi: 10.3386/w27812

Attanasio, O., Boneva, T., & Rauh, C. (2022). Parental beliefs about returns to different types of investments in school children. *Journal of Human Resources*, *57*(6), 1789–1825.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2020). Estimating the production function for human capital: results from a randomized controlled trial in colombia. *American Economic Review*, *110*(1), 48–85.

Attanasio, O., Meghir, C., & Nix, E. (2020). Human capital development and parental investment in india. *The Review of Economic Studies*, *87*(6), 2511–2541.

Barcellos, S. H., Carvalho, L. S., & Lleras-Muney, A. (2014). Child gender and parental investments in india: Are boys and girls treated differently? *American Economic Journal: Applied Economics*, *6*(1), 157–89.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics*, *3*(2), 167–95.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, *81*(2, Part 2), S279–S288.

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, *84*(4, Part 2), S143–S162.

Behrman, J. R. (1988). Nutrition, health, birth order and seasonality: Intrahousehold allocation among children in rural india. *Journal of Development Economics*, *28*(1), 43–62.

Behrman, J. R. (1997). Intrahousehold distribution and the family. *Handbook of population and family economics*, *1*, 125–187.

Behrman, J. R., Pollak, R. A., & Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, *90*(1), 52–73.

Behrman, J. R., & Rosenzweig, M. R. (2004). Parental allocations to children: New evidence on bequest differences among siblings. *Review of Economics and Statistics*, *86*(2), 637–640.

Berry, J., Dizon-Ross, R., & Jagnani, M. (2020). *Not playing favorites: An experiment on parental fairness preferences* (Tech. Rep.). National Bureau of Economic Research.

Bhalotra, S. R., Delavande, A., Gilabert, P., & Maselko, J. (2020). Maternal investments in children: The role of expected effort and returns.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, *120*(2), 669–700.

Boneva, T., & Rauh, C. (2016). Human capital production and parental beliefs.

Boneva, T., & Rauh, C. (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns.

Boneva, T., & Rauh, C. (2018). Parental beliefs about returns to educational investments - the later the better? *Journal of the European Economic Association*, *16*(6), 1669–1711.

Bradley, R. H., Caldwell, B. M., Rock, S. L., Hamrick, H. M., & Harris, P. (1988). Home observation for measurement of the environment: Development of a home inventory for use with families having children 6 to 10 years old. *Contemporary Educational Psychology*, *13*(1), 58–71.

Brito, N. H., & Noble, K. G. (2014). Socioeconomic status and structural brain development. *Frontiers in neuroscience*, *8*, 276.

Brown, C., Calvi, R., & Penglase, J. (2020). Sharing the pie: Undernutrition, intra-household allocation, and poverty.

Brown, C., Ravallion, M., & Van De Walle, D. (2017). *Are poor individuals mainly found in poor households? evidence using nutrition data for africa.* The World Bank.

Cáceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources*, *41*(4), 738–754.

Calvi, R. (2020). Why are older women missing in india? the age profile of bargaining power and poverty. *Journal of Political Economy*, *128*(7), 2453–2501.

Canning, D., & Schultz, T. P. (2012). The economic consequences of reproductive health and family planning. *The Lancet*, *380*(9837), 165–171.

Caplin, A. (2016). Measuring and modeling attention. *Annual Review of Economics*, *8*, 379–403.

Caplin, A. (2021). *Economic data engineering* (Tech. Rep.). National Bureau of Economic Research.

Caplin, A., & Martin, D. (2021). Comparison of decisions under unknown experiments. *Journal of Political Economy*, *129*(11), 3185–3205.

Chao, F., Guilmoto, C. Z., KC, S., & Ombao, H. (2020). Probabilistic projection of the sex ratio at birth and missing female births by state and union territory in india. *PloS one*, *15*(8), e0236673.

Chiappori, P.-A., & Meghir, C. (2015). Intrahousehold inequality. In *Handbook of income distribution* (Vol. 2, pp. 1369–1418). Elsevier.

Conti, G., Giannola, M., & Toppeta, A. (2022, Dec). *Parental beliefs, perceived health risks, and time investment in children: Evidence from covid-19* (IZA Discussion Paper No. 15765). Bonn: Institute of Labor Economics (IZA). Retrieved from https://covid-19.iza.org/publications/dp15765/

Cowell, F. (2011). *Measuring inequality*. Oxford University Press.

Cunha, F., Elo, I., & Culhane, J. (2013). *Eliciting maternal expectations about the technology of cognitive skill formation* (Tech. Rep.). National Bureau of Economic Research.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Currie, J., & Almond, D. (2011). Human capital development before age five. In *Handbook of labor economics* (Vol. 4, pp. 1315–1486). Elsevier.

Datar, A., Kilburn, M. R., & Loughran, D. S. (2010). Endowments and parental investments in infancy and early childhood. *Demography*, *47*(1), 145–162.

Deaton, A. (1989). Looking for boy-girl discrimination in household expenditure data. *The World Bank Economic Review*, *3*(1), 1–15.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. The World Bank.

Delavande, A. (2014). Probabilistic expectations in developing countries. *Annu. Rev. Econ.*, *6*(1), 1–20.

Delavande, A., & Kohler, H.-P. (2009). Subjective expectations in the context of hiv/aids in malawi. *Demographic research*, *20*, 817.

Delavande, A., & Zafar, B. (2019). University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints. *Journal of Political Economy*, *127*(5), 2343–2393.

Desai, S., Dubey, A., Joshi, B., Sen, M., Shariff, A., Vanneman, R., & Codebook, H. (2005). India human development survey (ihds). *NCAER, India and Inter-university Consortium for Political and Social Research, MI, USA*.

Desai, S., & Vanneman, R. (2015). *India human development survey-ii (ihds-ii), 2011-12.* Inter-university Consortium for Political and Social Research Ann Arbor, MI.

Di Tella, R., Galiani, S., & Schargrodsky, E. (2007). The formation of beliefs: evidence from the allocation of land titles to squatters. *The Quarterly Journal of Economics*, *122*(1), 209–241.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, *109*(8), 2728–65.

Doss, C., Fahle, E. M., Loeb, S., & York, B. N. (2019). More than just a nudge supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, *54*(3), 567–603.

Duflo, E. (2005). Gender equality in development. *BREAD Policy Paper*, *11*(4).

Dunbar, G. R., Lewbel, A., & Pendakur, K. (2013). Children's resources in collective households: identification, estimation, and an application to child poverty in malawi. *American Economic Review*, *103*(1), 438–71.

Duncan, G., Kalil, A., Mogstad, M., & Rege, M. (2022). Investing in early childhood development in preschool and at home.

Dupas, P. (2011). Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. *American Economic Journal: Applied Economics*, *3*(1), 1–34.

Fiszbein, A., & Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty.* World Bank Publications.

Fitzsimons, E., Malde, B., Mesnard, A., & Vera-Hernández, M. (2016). Nutrition, information and household behavior: Experimental evidence from malawi. *Journal of Development Economics*, *122*, 113–126.

Frankenberg, E., Karoly, L. A., Gertler, P., Achmad, S., Agung, I., Hatmadji, S. H., & Sudharto, P. (1995). The 1993 indonesian family life survey: Overview and field report.

Glewwe, P., & Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the economics of education* (Vol. 5, pp. 653–743). Elsevier.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy*, *87*(5, Part 2), S37–S64.

Gupta, M. D. (1987). Selective discrimination against female children in rural punjab, india. *Population and development review*, 77–100.

Haddad, L., & Kanbur, R. (1990). How serious is the neglect of intra-household inequality? *Economic Journal*, *100*(402), 866–81.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of political economy*, *100*(1), 84–117.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689–733.

Hjort, J., Moreira, D., Rao, G., & Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 brazilian municipalities. *American Economic Review*, *111*(5), 1442–80.

IIPS. (2001). National family health survey (NFHS-2), india, 1998-99: Orissa. *International Institute for Population Sciences, and Macro International. Mumbai: IIPS.*.

IIPS. (2008). National family health survey (NFHS-3), india, 2005-06: Orissa. *International Institute for Population Sciences and Macro International. Mumbai: IIPS.*.

ILO. (2018). India wage report: Wage policies for decent work and inclusive growth. *India: International Labour Organization*.

Indian Census. (2011).

Jayachandran, S. (2017). Fertility decline and missing women. *American Economic Journal: Applied Economics*, *9*(1), 118–39.

Jayachandran, S., & Kuziemko, I. (2011). Why do mothers breastfeed girls less than boys? evidence and implications for child health in india. *The Quarterly journal of economics*, *126*(3), 1485–1538.

Jayachandran, S., & Pande, R. (2017). Why are indian children so short? the role of birth order and son preference. *American Economic Review*, *107*(9), 2600–2629.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548.

Kaufmann, K. M. (2014). Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns. *Quantitative Economics*, *5*(3), 583–630.

Leight, J. (2017). Sibling rivalry: Endowment and intrahousehold allocation in gansu province, china. *Economic Development and Cultural Change*, *65*(3), 457–493.

Liebman, J. B., & Luttmer, E. F. (2015). Would people behave differently if they better understood social security? evidence from a field experiment. *American Economic Journal: Economic Policy*, *7*(1), 275–99.

Lise, J., & Seitz, S. (2011). Consumption inequality and intra-household allocations. *The Review of Economic Studies*, *78*(1), 328–355.

List, J. A., Pernaudet, J., & Suskind, D. L. (2021). Shifting parental beliefs about child development to foster parental investments and improve school readiness outcomes. *Nature communications*, *12*(1), 1–10.

Lochner, L., & Monge-Naranjo, A. (2012). Credit constraints in education. *Annu. Rev. Econ.*, *4*(1), 225–256.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, *72*(5), 1329–1376.

Mas, A., & Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, *107*(12), 3722–59.

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2019). Using behavioral insights to increase parental engagement the parents and children together intervention. *Journal of Human Resources*, *54*(4), 900–925.

Mogstad, M., & Wiswall, M. (2012). Instrumental variables estimation with partially missing instruments. *Economics Letters*, *114*(2), 186–189.

Mogstad, M., & Wiswall, M. (2016). Testing the quantity–quality model of fertility: Estimation using unrestricted family size models. *Quantitative Economics*, *7*(1), 157–192.

Padhi, S. (2001). Infant and child survival in orissa: An analysis with nfhs data. *Economic and Political Weekly*, 3316–3326.

Pitt, M. M., Rosenzweig, M. R., & Hassan, M. N. (1990). Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries. *The American Economic Review*, 1139–1156.

Ravallion, M. (2015). *The economics of poverty: History, measurement, and policy.* Oxford University Press.

Reserve Bank of India . (2017). (Tech. Rep.).

Rosenzweig, M. R., & Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of political economy*, *91*(5), 723–746.

Rosenzweig, M. R., & Wolpin, K. I. (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: journal of the Econometric Society*, 227–240.

Rosenzweig, M. R., & Wolpin, K. I. (1988). Heterogeneity, intrafamily distribution, and child health. *Journal of human resources*, 437–461.

Rosenzweig, M. R., & Zhang, J. (2009). Do population control policies induce more human capital investment? twins, birth weight and china's "one-child" policy. *The Review of Economic Studies*, *76*(3), 1149–1174.

Rubalcava, L., & Teruel, G. (2013). Mexican family life survey, third wave. *Retrieved (www. ennvih-mxfls. org)*.

Soares, R. R. (2005). Mortality reductions, educational attainment, and fertility choice. *American Economic Review*, *95*(3), 580–601.

Solis, A. (2017). Credit access and college enrollment. *Journal of Political Economy*, *125*(2), 562–622.

Stantcheva, S. (2022). *How to run surveys: A guide to creating your own identifying variation and revealing the invisible* (Tech. Rep.). National Bureau of Economic Research.

Stinebrickner, R., & Stinebrickner, T. R. (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, *81*(1), 426–472.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124–1131.

United Nations. (2019). World population prospects 2019, online edition. rev. 1. *New York (US): United Nations Department for Economic and Social Affairs, Population Division.*

WHO. (2019). *World health statistics overview 2019: monitoring health for the sdgs, sustainable development goals* (Tech. Rep.). World Health Organization.

Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, *82*(2), 791–824.

Wiswall, M., & Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, *133*(1), 457–507.

Yadav, A. K., Anand, A., Singh, R. A., & Jungari, S. (2020). Sex composition and son preference in india and major states: Evidence from the national family health survey-4 (2015-16). *Clinical Epidemiology and Global Health*.

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal*, *125*(588), F347–F371.

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), 537–566.

# APPENDIX

## A Appendix Tables and Figures

| | Perceived earnings at age 30 (log) | | | |
|---|---|---|---|---|
| | Full sample | | Control group | |
| | (1) | (2) | (3) | (4) |
| High ability | 0.897*** | 0.837*** | 0.878*** | 0.818*** |
| | (0.046) | (0.046) | (0.054) | (0.054) |
| High Investment | 0.234*** | 0.175*** | 0.234*** | 0.175*** |
| | (0.018) | (0.023) | (0.018) | (0.023) |
| Treatment | -0.024 | -0.034 | | |
| | (0.064) | (0.065) | | |
| High ability × Treatment | 0.031 | 0.049 | | |
| | (0.055) | (0.058) | | |
| High investment × Treatment | 0.014 | 0.033 | | |
| | (0.025) | (0.037) | | |
| Ability × Investment | | 0.119*** | | 0.119*** |
| | | (0.027) | | (0.027) |
| High ability × High investment × Treatment | | -0.038 | | |
| | | (0.043) | | |
| Boy | 0.154*** | 0.154*** | 0.119 | 0.119 |
| | (0.056) | (0.056) | (0.076) | (0.076) |
| Mean outcome | 29676 | 29676 | 29428 | 29428 |
| $R^2$ | 0.364 | 0.365 | 0.356 | 0.358 |
| Observations | 2480 | 2480 | 1296 | 1296 |

*Notes:* This table presents analogous coefficients and standard errors to those presented in Table 2. In columns 1 and 2 all the main regressors are interacted with RCT treatment status. In column 3 and 4 the estimation sample is restricted to household in the control group. Because treatment status is allocated at the respondent level, regressions do not control for family fixed effects. The relevant comparison for columns 1 and 3 is column 2 from Table 2. The relevant comparison for columns 2 and 4 is column 3 of Table 2. The number of observations is smaller because not all participants to the survey experiment participated to the original RCT. As explained in the text in larger slums neighbours of randomly selected households from the original experimental sample were also interviewed. Robust standard errors clustered at the family level are reported in brackets. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.1: Effect of RCT Treatment Status on Perceived Production Function

| | Share of resources to child $H$ | | | |
|---|---|---|---|---|
| | Full sample | | Control group | |
| | (1) | (2) | (3) | (4) |
| Difference in ability | 0.077*** | 0.109*** | 0.077*** | 0.109*** |
| | (0.010) | (0.015) | (0.010) | (0.015) |
| Treatment | -0.004 | -0.006 | | |
| | (0.013) | (0.019) | | |
| High ability × Treatment | 0.001 | -0.014 | | |
| | (0.014) | (0.021) | | |
| High resource | | 0.033*** | | 0.033*** |
| | | (0.011) | | (0.011) |
| Difference in ability × High resources | | -0.064*** | | -0.064*** |
| | | (0.014) | | (0.014) |
| High resources × Treatment | | 0.005 | | |
| | | (0.016) | | |
| Difference in ability × High resources × Treatment | | 0.032 | | |
| | | (0.021) | | |
| Boy | 0.006 | 0.006 | 0.013 | 0.013 |
| | (0.011) | (0.011) | (0.016) | (0.016) |
| $R^2$ | 0.076 | 0.087 | 0.073 | 0.085 |
| Observations | 1240 | 1240 | 648 | 648 |

*Notes:* This table presents analogous coefficients and standard errors to those presented in Table 3. In columns 1 and 2 all the main regressors are interacted with RCT treatment status. In column 3 and 4 the estimation sample is restricted to household in the control group. Because treatment status is allocated at the respondent level, regressions do not control for family fixed effects. The relevant comparison for columns 1 and 3 is column 1 from Table 3. The relevant comparison for columns 2 and 4 is column 3 of Table 3 . The number of observations is smaller because not all participants to the survey experiment participated to the original RCT. As explained in the text in larger slums neighbours of randomly selected households from the original experimental sample were also interviewed. Robust standard errors clustered at the family level are reported in brackets. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.2: Effect of RCT Treatment Status on Allocation of Resources

| | Educational attainment (in years) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| High ability | 2.284*** | 2.283*** | 2.288*** | 2.288*** |
| | (0.052) | (0.061) | (0.062) | (0.057) |
| High Investment | 0.267*** | 0.267*** | 0.272*** | 0.272*** |
| | (0.021) | (0.021) | (0.027) | (0.029) |
| Boy | | -0.001 | -0.001 | |
| | | (0.063) | (0.063) | |
| Ability × Investment | | | -0.010 | -0.010 |
| | | | (0.035) | (0.037) |
| Observations | 3960 | 3960 | 3960 | 3960 |
| Mean outcome | 11.471 | 11.471 | 11.471 | 11.471 |
| $R^2$ | 0.542 | 0.542 | 0.542 | 0.775 |
| Fixed effects | | | | ✓ |

*Notes:* The outcome variable is educational attainment (in years) as perceived by the respondent. Columns 1 to 3 display the OLS results. Columns 4 further include family fixed effects. Robust standard errors clustered at the family level are reported in brackets. *High ability* is a dummy variable that takes value 1 if in scenario $j$ the child has an high academic ability, *High investment* is a dummy variable that takes value one if in scenario $j$ the level of investments is high, and *Boy* is a dummy variable equal to one if the child is a boy. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.3: Perceived Production Function (Educational Attainment)

| | Perceived earnings at age 30 (log) | | | | | |
|---|---|---|---|---|---|---|
| | Girls | | | Boys | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| High ability | 0.836*** | 0.776*** | 0.371*** | 0.826*** | 0.760*** | 0.438*** |
| | (0.030) | (0.035) | (0.056) | (0.029) | (0.033) | (0.058) |
| High Investment | 0.260*** | 0.200*** | 0.156*** | 0.245*** | 0.179*** | 0.137*** |
| | (0.015) | (0.023) | (0.022) | (0.015) | (0.021) | (0.022) |
| Ability × Investment | | 0.120*** | 0.114*** | | 0.133*** | 0.140*** |
| | | (0.028) | (0.027) | | (0.027) | (0.027) |
| Belief about child education | | | 0.174*** | | | 0.143*** |
| | | | (0.021) | | | (0.026) |
| Mean outcome | 27710 | 27710 | 27710 | 32539 | 32539 | 32539 |
| $R^2$ | 0.353 | 0.775 | 0.807 | 0.381 | 0.766 | 0.788 |
| Observations | 1984 | 1984 | 1984 | 1976 | 1976 | 1976 |
| Fixed effects | | ✓ | ✓ | | ✓ | ✓ |

*Notes:* The table report coefficients analogous to those presented in Table 2 by splitting the sample according to the gender of the two children. The first 3 columns report the results for the sample of respondent who saw one boy and one girl, while the remaining 3 columns report results for the sample who saw two boys. The outcome variable is log-earnings of the child at age 30 as perceived by the respondent. Columns 1 and 4 display the OLS results. Columns 2, 3, 5 and 6 further include family fixed effects. Robust standard errors clustered at the family level are reported in brackets. High ability is a dummy variable that takes value 1 if in scenario j the child has an high initial skill level, High investment is a dummy variable that takes value one if in scenario j the level of investments is high. Belief about child education is the educational attainment respondents believe the child will achieve in scenario $j$. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.4: Perceived Production Function by Gender

| | Educational attainment (in years) | | | |
|---|---|---|---|---|
| | Girls | | Boys | |
| | (1) | (2) | (3) | (4) |
| High ability | 2.347*** | 2.330*** | 2.220*** | 2.247*** |
| | (0.074) | (0.082) | (0.072) | (0.081) |
| High Investment | 0.269*** | 0.252*** | 0.266*** | 0.293*** |
| | (0.032) | (0.044) | (0.026) | (0.037) |
| Ability × Investment | | 0.034 | | -0.054 |
| | | (0.054) | | (0.051) |
| Mean outcome | 11 | 11 | 12 | 12 |
| $R^2$ | 0.556 | 0.771 | 0.528 | 0.780 |
| Observations | 1984 | 1984 | 1976 | 1976 |
| Fixed effects | | ✓ | | ✓ |

*Notes:* The table report coefficients analogous to those presented in Table A.3 by splitting the sample according to the gender of the two children. The first 2 columns report the results for the sample of respondent who saw one boy and one girl, while the remaining 2 columns report results for the sample who saw two boys. The outcome variable is educational attainment (in years) as perceived by the respondent. Columns 1 and 3 display the OLS results. Columns 2 and 4 further include family fixed effects. Robust standard errors clustered at the family level are reported in brackets. High ability is a dummy variable that takes value 1 if in scenario j the child has an high initial skill level, High investment is a dummy variable that takes value one if in scenario j the level of investments is high. Belief about child education is the educational attainment respondents believe the child will achieve in scenario $j$. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.5: Perceived Production Function by Gender (Educational Attainment)

| | Earnings | | | Education | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| High ability | 0.836*** | 0.776*** | 0.311*** | 2.347*** | 2.330*** |
| | (0.030) | (0.032) | (0.066) | (0.074) | (0.076) |
| High Investment | 0.260*** | 0.200*** | 0.150*** | 0.269*** | 0.252*** |
| | (0.015) | (0.021) | (0.021) | (0.032) | (0.041) |
| Boy | 0.176*** | 0.180*** | 0.101 | 0.127 | 0.105 |
| | (0.048) | (0.049) | (0.383) | (0.101) | (0.102) |
| Ability × Investment | | 0.120*** | 0.113*** | | 0.034 |
| | | (0.026) | (0.026) | | (0.051) |
| High ability × Boy | -0.009 | -0.015 | -0.011 | -0.127 | -0.083 |
| | (0.042) | (0.045) | (0.090) | (0.103) | (0.108) |
| High investment × Boy | -0.015 | -0.021 | -0.031 | -0.003 | 0.040 |
| | (0.021) | (0.029) | (0.030) | (0.042) | (0.054) |
| Ability × Investment x Boy | | 0.013 | 0.030 | | -0.088 |
| | | (0.036) | (0.036) | | (0.070) |
| Belief about child education | | | 0.200*** | | |
| | | | (0.025) | | |
| Belief about child education × Boy | | | 0.006 | | |
| | | | (0.037) | | |
| Mean outcome | 30120 | 30120 | 30120 | 11 | 11 |
| $R^2$ | 0.374 | 0.376 | 0.464 | 0.542 | 0.543 |
| Observations | 3960 | 3960 | 3960 | 3960 | 3960 |

*Notes:* The table report coefficients analogous to those presented in Table 2 where the main coefficients are interacted with an indicator for whether the respondent was randomised in a group that saw two boys or one boy and one girl. In the first 3 columns the outcome variable is log-earnings of the child at age 30 as perceived by the respondent. While in columns 4 and 5 the outcome variable is educational attainment (in years) as perceived by the respondent. Robust standard errors clustered at the family level are reported in brackets. High ability is a dummy variable that takes value 1 if in scenario j the child has an high initial skill level, High investment is a dummy variable that takes value one if in scenario j the level of investments is high. Belief about child education is the educational attainment respondents believe the child will achieve in scenario $j$. $^{*}$ denotes 10% significance, $^{**}$ denotes 5% significance, $^{***}$ denotes 1% significance.

Table A.6: Perceived Production Function by Gender - Fully interacted model

| | School fees | Uniforms | Textbooks | Stationary | Private tuition† |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Ability | 107.120 | 37.615** | 42.250* | 21.187 | 233.388** |
| | (110.122) | (16.891) | (21.864) | (16.758) | (92.177) |
| Child controls | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean outcome | 3316 | 522 | 681 | 850 | 3374 |
| Observations | 582 | 582 | 582 | 582 | 556 |

*Notes:* The outcome variable is the difference in investment between two children, in the outcome variable shown in the column header. Expenditure is measured in rupees and is reported at the yearly level unless otherwise specified. Child controls include gender and age. Robust standard errors clustered at the family level are reported in brackets. † Tuition expenditure was only collected at the monthly level. This is converted in yearly expenditure by multiplying by 10. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.7: Actual Parental Investments: Expenditure Category

| | Encourage to read | Conversations | Outings | Play | Homework | Discuss school |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ability | 0.058*** | 0.065*** | 0.038*** | -0.001 | 0.025** | 0.043*** |
| | (0.007) | (0.008) | (0.010) | (0.009) | (0.011) | (0.009) |
| Child controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean | 0.91 | 0.90 | 0.56 | 0.33 | 0.68 | 0.69 |
| Observations | 417 | 417 | 417 | 417 | 417 | 417 |

*Notes:* The outcome is the difference in investment between two children, in the variable shown in the column header. Child controls include gender and age. Robust standard errors clustered at the family level are reported in brackets. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.8: Actual Parental Investments: Time

| | Educational expenditure | | | Child work | | |
|---|---|---|---|---|---|---|
| | All | $\rho_L$ | $\rho_H$ | All | $\rho_L$ | $\rho_H$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ability | 278.001 | 178.261 | 713.596** | -0.042*** | -0.031** | -0.075*** |
| | (187.180) | (232.077) | (342.879) | (0.011) | (0.013) | (0.021) |
| Child controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 434 | 214 | 220 | 370 | 185 | 185 |
| | Private school | | | Time index | | |
| | All | $\rho_L$ | $\rho_H$ | All | $\rho_L$ | $\rho_H$ |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| Ability | 0.005 | 0.004 | 0.009 | 0.217*** | 0.185*** | 0.275*** |
| | (0.007) | (0.008) | (0.012) | (0.023) | (0.031) | (0.037) |
| Child controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 421 | 203 | 218 | 285 | 140 | 145 |

*Notes:* The outcome is the difference in investment between two children, as measured by the outcome in the column header. Families in the treatment group are dropped from the estimation sample. All means all households, $\rho_L$ means higher inequality aversion (above the median), and $\rho_H$ means lower inequality aversion. Child controls include gender and age. Robust standard errors clustered at the family level are reported in brackets. † in INR. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.9: Actual Parental Investments: Dropping Treatment Households

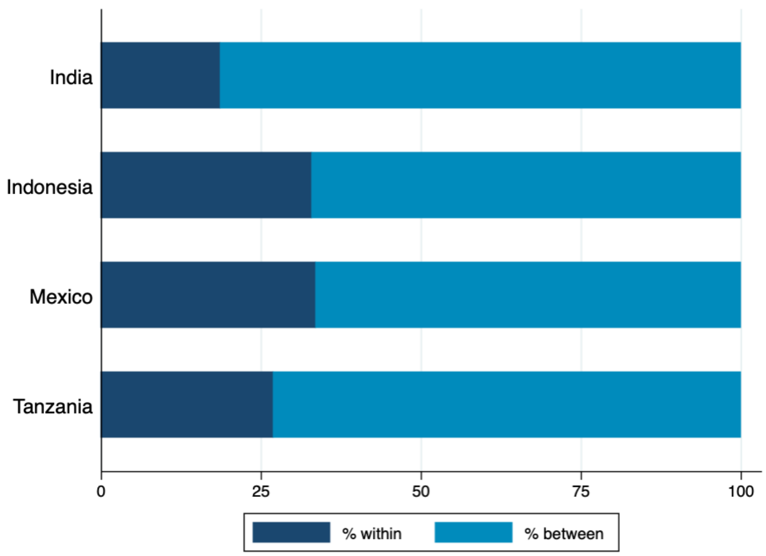



Figure A.1: Inequality in Child Human Capital (Test Scores)

*Notes:* This figure plots the within-household and between-households component of the Mean Log Deviation (MLD) measure of inequality. The outcome variable is test scores. I use an age-standardized z-score, where the reference group consists of children in the same country and age. Thus coefficients are expressed in standard deviations units. Each bar represents a different country. Source: Indian Human Development Survey (Desai et al. (2005), Desai & Vanneman (2015)), Mexican Family Life Survey (Rubalcava & Teruel (2013)), Indonesian Family Life Survey (Frankenberg et al. (1995)), Uwezo initiative for Tanzania.

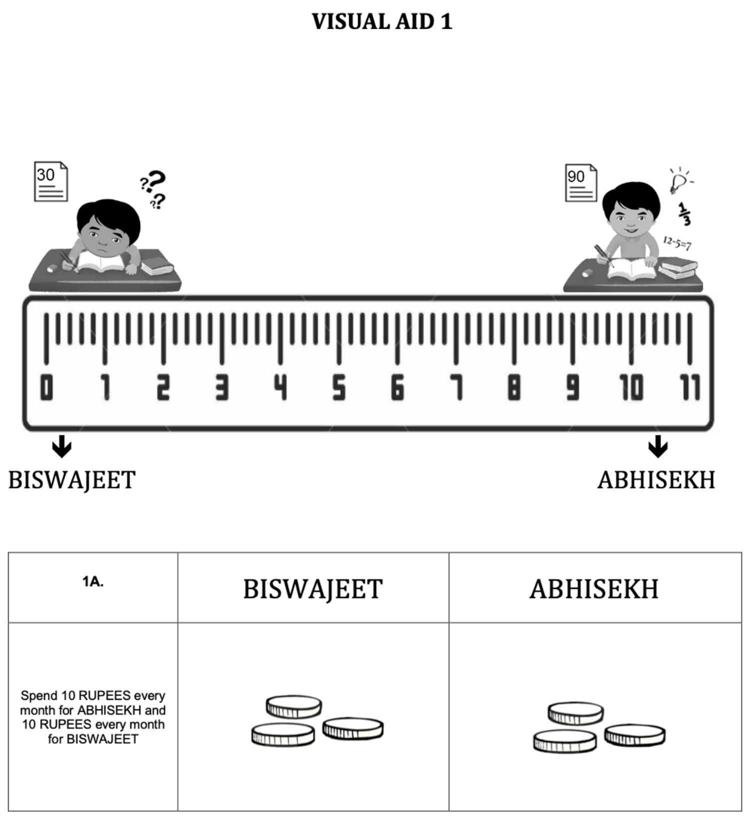
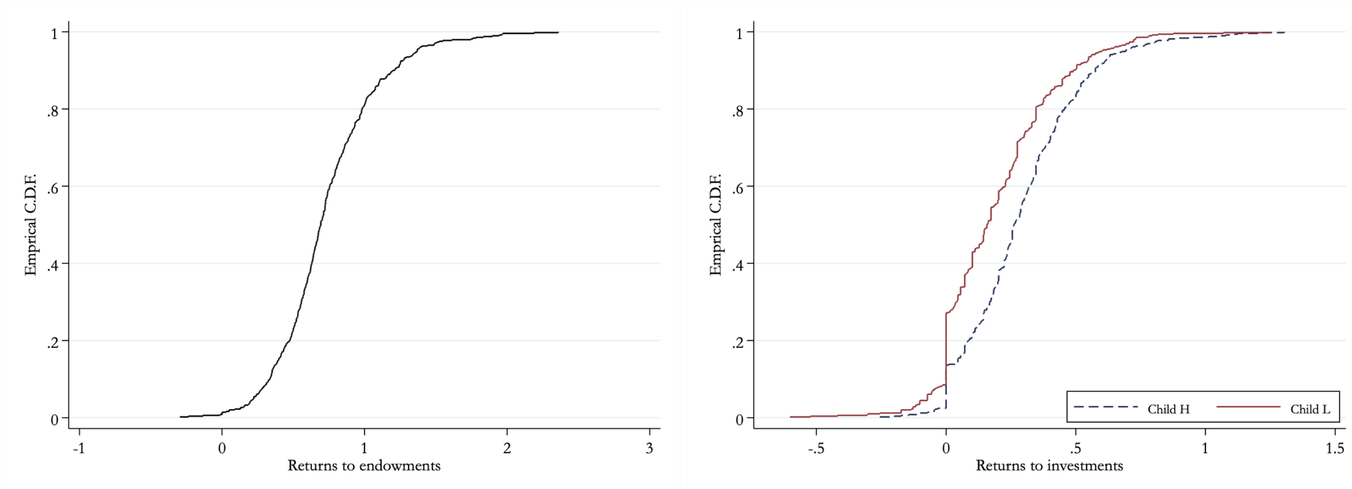



Figure A.2: Visual Aid

*Notes:* This figure shows an example of visual aid used to elicit parental beliefs about the human capital production function. Child baseline ability was described with the help of the ruler at the top of the figure. Parental investments where described using the coins at the bottom of the figure. In the example reported here one child is described has having a low baseline ability, while the other child as having a high baseline ability, and the level of investment in each child is low.




(A) Returns to child ability

(B) Returns to investments

Figure A.3: Heterogeneity in Perceived Returns

*Notes:* This figure plots the empirical CDF of individual perceived returns. Panel A shows the CDF for the perceived returns to ability, while panel B shows the CDF for the perceived return to investment. Panel B shows two CDFs. The solid one is for a child with low baseline ability, while the dashed one is for a child with high baseline ability.

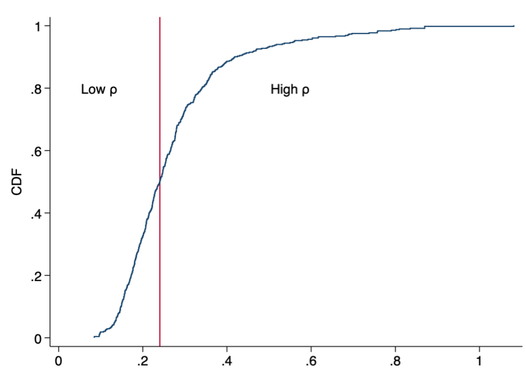
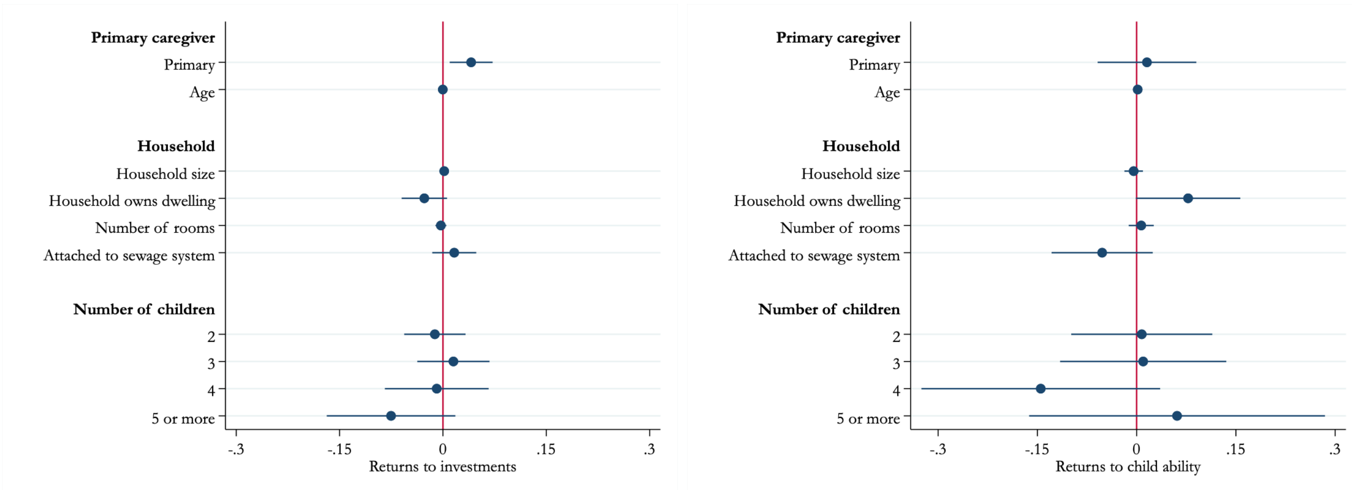



Figure A.4: Heterogeneity in Preferences

*Notes:* The figure plots the empirical CDF of parental preferences for intra-household inequality. The vertical line represents the median value of $\rho$ in the sample. Low $\rho$ households have greater concerns for intra-household inequality in child outcomes.




(A) Returns to investment

(B) Returns to child ability

Figure A.5: Correlations between perceived returns and observable characteristics

*Notes:* The figure plots the coefficients of a regression of individual perceived returns to investments (Panel A) and child ability (Panel B) from Section 6.1 on observable characteristics.

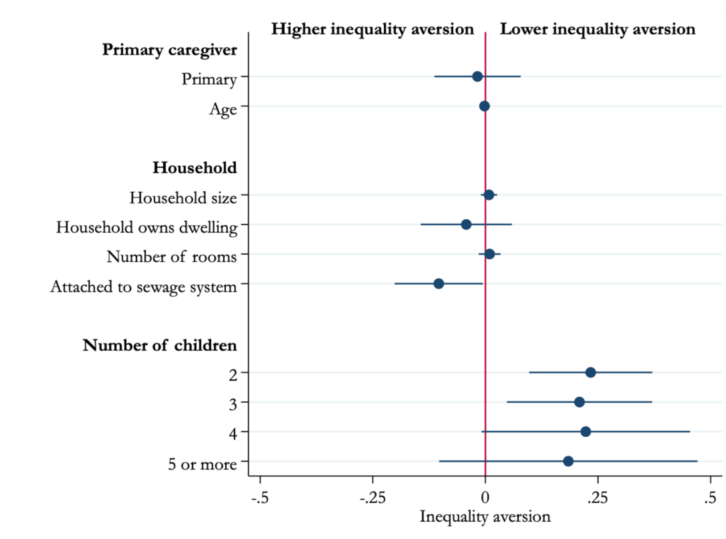



Figure A.6: Correlations between parental preferences and observable characteristics

*Notes:* The figure plots the coefficients of a regression of parental preferences for intra-household inequality from Section 6.2 on observable characteristics.

# B Model and Derivations

## B.1 Discussion of the model

This section discusses model assumptions and extensions.

**Fertility choices.** One assumption in the model is that parents choose child educational investments conditional on an exogenously given family size $n$. The theoretical framework can be easily extended to allow parents to choose fertility endogenous. To do so, assume that parents first decide sequentially on the number of children they have. Once the fertility spell is concluded, they decide how to allocate educational investments. The model can be solved backwards, and implies an optimal stopping problem. One can show that in each period parents compare the utility from having $n$ children with the expected utility of having $n+1$ children. They will stop when the former is greater than the latter (a formal derivation of the optimal stopping rule is available upon request). The optimal stopping rule describing fertility behaviour also depends on parental preferences for intra-household inequality (the parameter $\rho$). In particular, the model implies an endogenous fertility response to child ability so that parents are more likely to increase fertility after giving birth to child with a low $A_i$. Importantly, the optimal allocation rule in (7) is not affected by the fertility decision, so that the results derived in the previous section are still valid when allowing for endogenous fertility. If anything, those results are reinforced by the fact that, because of the optimal stopping rule, low ability children are more likely to belong to larger families, resulting in them having more siblings and thus facing more competition over limited resources. Using data from the Indian National Family and Health survey, I test and find empirical support for the prediction that parents are more likely to increase fertility after giving birth to child with a low $A_i$ (the results are available upon request). This prediction is also consistent with the demographic transition literature, which shows that reductions in child mortality are associated with a decline in fertility (Soares (2005)), and with a public health literature documenting that improvements is health at birth are associated with reductions in maternal fertility (Canning & Schultz (2012)).

**The Quantity-Quality trade-off.** When parents reinforce ability differences, the model implies the existence of a negative relation between family size and average child quality (the Quantity-Quality trade off), even if the maximum level of human capital stays constant as family size increases. This suggests that when parents target their investments to their children's ability, an increase in family size can differentially affect children living in the same family. Because of allocation of resources that take place within the household, changes in family size will have asymmetric effects on different children, so that average treatment effects might be misleading. In particular, while higher achieving children are less affected by variations in family size, the human capital of low achieving children sharply declines as family size increases because of less per-capita resources and more competition between siblings. This heterogeneous effect of family size on child outcomes could potentially explain why the empirical findings in the Quantity-Quality literature are mixed, with some studies finding evidence in favour of a trade-off (Rosenzweig & Wolpin (1980); Hanushek (1992); Rosenzweig & Zhang (2009); Mogstad & Wiswall (2016)), while other against (Black, Devereux, & Salvanes (2005); Angrist, Lavy, & Schlosser (2010); Cáceres-Delpiano (2006)). What the model suggests is that family size *per se* might have little effect on child human capital, what matters for child outcomes is the effect that family size has on household resources, combined with parental allocative decisions.

**Child specific preferences.** By including child-specific weights in the utility function (1), the model is general enough to incorporate preferences for some specific children or some specific characteristics of the child that might be important in some contexts. For instance, In India these is a larger literature suggesting that parents might a preference for sons over daughters (Gupta (1987); Jayachandran (2017)). This gender preference is particularly strong in for some parts of India – particularly in the North-West – and significantly less pronounced in other states (Jayachandran & Pande (2017); Yadav, Anand, Singh, & Jungari (2020)).

### B.2 Close form solution for investments

This sections derives closed form solutions for investments. Maximizing (1) subject to (2) and (6) one can get to the following closed form solution for investments in child $i$:

$$I_i^* = y \frac{c_i^{\frac{1}{1-b\rho}} \times A_i^{\frac{a\rho}{1-b\rho}}}{\sum_{j=1}^{n} c_j^{\frac{1}{1-b\rho}} \times A_j^{\frac{a\rho}{1-b\rho}}} \tag{B.1}$$

Computing the ratio of $I_i^*$ to $I_j^*$ and taking the log delivers equation (7).

### B.3 Derivation of equation (13)

Starting from equation (12):

$$\gamma = \frac{a\rho}{1-b\rho} \implies \gamma(1-b\rho) = a\rho \implies \gamma = a\rho + b\gamma\rho \tag{B.2}$$

Solving for $\rho$:

$$\rho(a+b\gamma) = \gamma \implies \rho = \frac{\gamma}{a+b\gamma} \tag{B.3}$$

Multiplying and diving the right hand side term by $a$:

$$\rho = \frac{1}{a} \times \frac{a\gamma}{a+b\gamma} \implies \rho = \frac{1}{a} \times \left[\frac{a+b\gamma}{a\gamma}\right]^{-1} \tag{B.4}$$

From which:

$$\rho = \frac{1}{a} \times \left[\frac{1}{\gamma} + \frac{b}{a}\right]^{-1} \tag{B.5}$$

# ONLINE APPENDIX (NOT FOR PUBLICATION)
# Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Survey Experiment
Michele Giannola

## A Mean Log Deviation Measure of Inequality

Figure 1 and Figure 2 use the Mean Log Deviation Measure of Inequality (MLD) to decompose overall inequality in a within-household and between-households components. The MLD can be expressed as:

$$MLD = \frac{1}{N} \sum_{i} ln \frac{\bar{y}}{y_i} \tag{A.6}$$

where $y_i$ is individual outcome, $\bar{y}$ is average outcome among all individuals, and $N$ is the total number of individuals. This measure can be decomposed into a within group and between groups components as follows:

$$MLD = \sum_{j} \frac{N_j}{N} MLD_j + \sum_{j} \frac{N_j}{N} ln \frac{\bar{y}}{\bar{y_j}} \tag{A.7}$$

where $N_j$ is the total size of group $j$, $MLD_j$ is the mean log deviation measure of inequality in group $j$ and $\bar{y_j}$ is the average outcome among all individuals in group $j$. The first term in the within-group component and the second the between-groups component (see Cowell (2011) for a formal derivation of this expression).

## B Robustness for Figure 2

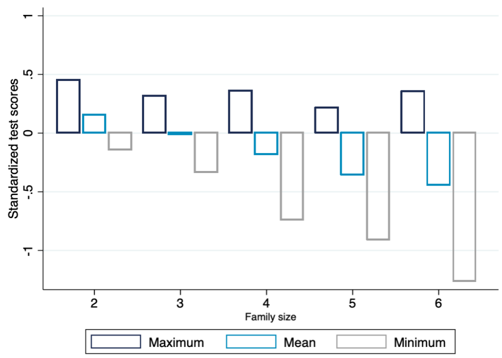
This section provides several robustness checks for the relation between fertility and the distribution of human capital in the family shown in Figure 2.

- Figure B.1 shows the relation between family size and the distribution of child quality using age standardized test scores as measure of quality. Each sub-plot represents a different country. The figure shows that the relation in Figure 2 holds across countries and is robust to the definition of human capital used.

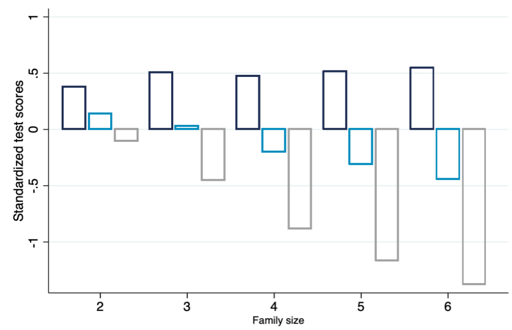
- Table B.1 report the regression results using age standardized test scores as measure of quality. In the table, I report the results of separate regression for the mean (columns 1 to 4), the maximum (columns 5 to 8) and the minimum (column 9 to 12). Columns 1, 5 and 9 include a linear indicator for family size. Columns 2, 6, and 10 include indicators for family size (top coded at size 6). Columns 3, 7 and 11 further control for birth order effects (top coded at birth order 6). Finally, columns 4, 8, and 12 include controls for mother and family background characteristics. All regressions control for child gender and age. The Table shows that the results are not driven by child background characteristics. The preferred specifications in columns 4, 8 and 12 (that control for child and maternal background characteristics) reveal a clear negative gradient in quality of the lowest achieving child in the family (column 12), and a shallow gradient in the quality of the highest achiever (column 8). Indeed, none of the family size dummies in column 8 is statistically different from zero and there is no clear patterns in the coefficients with some being negative while other positive. Comparing the coefficients in column 2 and 3, we can also infer that there is a negative birth order gradient in child human capital (the birth order dummies have been omitted to avoid clutter): once birth order is controlled for, the effect of family size on child outcomes becomes smaller in magnitude.

- Table B.2 report similar regression results as in Table B.1, but restricting the sample to women who have completed their fertility spell as identified in Jayachandran & Pande (2017). The outcome variable is age standardized test scores. Regressions control for birth order dummies, (top coded at birth order 6), child gender, child age and mother characteristics. These include maternal education dummies and location fixed-effects. All regressions control for child gender and age. The Table confirms the results from Table B.1: there is a strong negative gradient in the minimum and a shallow gradient in the maximum.

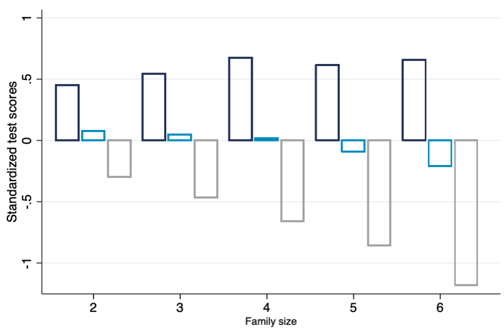
- Table B.3 reports the IV results using years of schooling as measure of human capital. Family size is instrumented using twin birth as an instrument for total family size. In the table, I report separate regressions for the mean, the maximum and the minimum. Panel A reports the results for India, while panel B reports the results for the other developing countries shown in Figure 2. I follow Angrist, Lavy, & Schlosser (2010) and report the results for the parity-pooled estimates to gain statistical power (i.e. I pool the 2+,3+, 4+ and 5+samples including first born in families with at least two births, first and second born in families with at least 3 births etc...). I account for missing instruments using the procedure introduced in Mogstad & Wiswall (2012). The Table confirms the results from Table B.1. There is a negative and significant effect of family size on the human capital of the lowest achieving child in the family, and a null effect on the human capital of the highest achieving child.



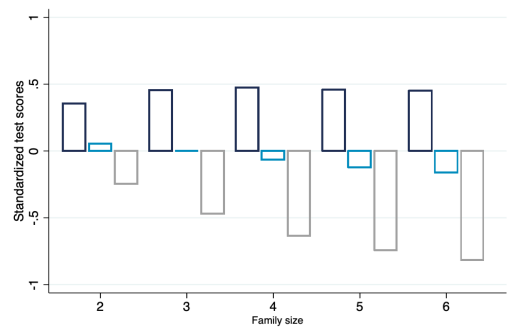

(A) India




(B) Mexico




(C) Indonesia




(D) Tanzania

Figure B.1: Fertility and Inequality in Child Human Capital (Test Scores)

*Notes:* The figure shows the relationship between family size and the mean (light blue), the maximum (dark blue) and the minimum (grey) levels of human capital within the household. This figure is constructed as follows. For each family in the sample, I compute the maximum, minimum and mean levels of human capital achieved by children in that family. For each level of fertility, I then average across families. The outcome variable is test scores. I use an age-standardized z-score, where the reference group consists of children in the same country and of the same age. Thus coefficients are expressed in standard deviations units. Source: Indian Human Development Survey (Desai et al. (2005), Desai & Vanneman (2015)), Mexican Family Life Survey Rubalcava & Teruel (2013), Indonesian Family Life Survey Frankenberg et al. (1995), Uwezo initiative for Tanzania.

| | Mean | | | | Maximum | | | | Minimum | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Linear family size | -0.149*** | | | | -0.063*** | | | | -0.266*** | | | |
| | (0.011) | | | | (0.015) | | | | (0.017) | | | |
| Family dummies | | | | | | | | | | | | |
| 3 children | | -0.146*** | -0.110*** | -0.002 | | -0.134*** | -0.152*** | -0.053 | | -0.192*** | -0.112*** | 0.021 |
| | | (0.025) | (0.027) | (0.026) | | (0.030) | (0.033) | (0.032) | | (0.033) | (0.036) | (0.035) |
| 4 children | | -0.296*** | -0.238*** | -0.015 | | -0.091* | -0.096 | 0.113* | | -0.597*** | -0.498*** | -0.237*** |
| | | (0.039) | (0.043) | (0.042) | | (0.055) | (0.061) | (0.060) | | (0.060) | (0.068) | (0.066) |
| 5 children | | -0.459*** | -0.380*** | -0.138*** | | -0.235*** | -0.249*** | -0.014 | | -0.767*** | -0.622*** | -0.355*** |
| | | (0.042) | (0.050) | (0.049) | | (0.060) | (0.071) | (0.070) | | (0.066) | (0.081) | (0.077) |
| 6 or more children | | -0.577*** | -0.427*** | -0.162* | | -0.100 | -0.167 | 0.086 | | -1.121*** | -0.825*** | -0.508*** |
| | | (0.067) | (0.087) | (0.084) | | (0.112) | (0.140) | (0.136) | | (0.123) | (0.167) | (0.157) |
| F-test | | 48.665 | 18.834 | 2.757 | | 6.829 | 6.597 | 2.623 | | 65.455 | 23.835 | 10.004 |
| p-value† | | 0.000 | 0.000 | 0.026 | | 0.000 | 0.000 | 0.033 | | 0.000 | 0.000 | 0.000 |
| Observations | 6315 | 6315 | 6315 | 6291 | 3069 | 3069 | 3069 | 3057 | 3069 | 3069 | 3069 | 3057 |
| Birth order dummies | | | ✓ | ✓ | | | ✓ | ✓ | | | ✓ | ✓ |
| Mother characteristics | | | | ✓ | | | | ✓ | | | | ✓ |

*Notes*: The outcome variables are standardized test scores. Columns 1 to 4 display the results for average levels of human capital, pooling all children together. Columns 5 to 8 display the results for the maximum (i.e. one child per family). Columns 9 to 12 display the results for the minimum (i.e. one child per family). Columns 1, 5 and 9 includes a linear indicator of family size. Column 2, 6 and 10 includes total fertility dummies, top-coded at 6 children. Column 3, 7 and 11 includes total fertility dummies (top-coded at 6 children) and birth order dummies (top coded at birth order 6). Columns 4, 8 and 12 includes total fertility dummies (top-coded at 6 children), birth order dummies (top coded at birth order 6) and mother characteristics. This include maternal education dummies and location fixed-effects. All regressions control for child gender and child age. Standard errors are reported in brackets. † p-value of an F-test on the joint significance of the family size dummies. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table B.1: Effect of Fertility on the Distribution of Human Capital in the Family

| | Mean | Maximum | Minimum |
|---|---|---|---|
| | (1) | (2) | (3) |
| Family size dummies | | | |
| 3 children | 0.050 | -0.015 | 0.090 |
| | (0.049) | (0.062) | (0.070) |
| 4 children | 0.006 | 0.116 | -0.240* |
| | (0.087) | (0.121) | (0.135) |
| 5 children | -0.082 | 0.198 | -0.376** |
| | (0.090) | (0.132) | (0.150) |
| 6 children or more | -0.268** | 0.056 | -0.752*** |
| | (0.130) | (0.198) | (0.292) |
| F-test | 2.06 | 0.96 | 4.61 |
| p-value† | 0.08 | 0.43 | 0.00 |
| Observations | 3595 | 1109 | 1111 |

*Notes:* The outcome variables are standardized test scores. The sample used in these regression is the same as that used in Jayachandran & Pande (2017). All regressions include total fertility dummies (top-coded at 6 children), birth order dummies (top coded at birth order 6) and mother characteristics. These include maternal education dummies and location fixed-effects. All regressions control for child gender and child age. Standard errors are reported in brackets. † p-value of an F-test on the joint significance of the family size dummies. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table B.2: Effect of Fertility on the Distribution of Human Capital in the Family - Completed Fertility Sample

| | OLS | | | IV | | |
|---|---|---|---|---|---|---|
| | Mean | Maximum | Minimum | Mean | Maximum | Minimum |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Panel A: India* | | | | | | |
| Linear family size | -0.081*** | 0.003 | -0.163*** | -0.053* | -0.000 | -0.156*** |
| | (0.001) | (0.002) | (0.002) | (0.031) | (0.024) | (0.025) |
| Observations | 366031 | 160199 | 153066 | 366031 | 160199 | 153066 |
| *Panel B: Developing countries* | | | | | | |
| Linear family size | -0.043*** | 0.025*** | -0.112*** | 0.004 | 0.020 | -0.050*** |
| | (0.001) | (0.001) | (0.001) | (0.028) | (0.015) | (0.017) |
| Observations | 393215 | 177587 | 169086 | 393215 | 177587 | 169086 |

*Notes:* The outcome variable is years of schooling (age-standardized z-score). The reference group consists of children in the same country and birth cohort. In each regression we pool the 2+, 3+, 4+ and 5+ samples together (as defined in Angrist, Lavy, & Schlosser (2010)). Columns 1 to 3 display the OLS results and columns 4 to 6 display the IV results. All regressions control for child gender, child age, child age squared, mother year of birth, household wealth index and maternal education. Standard errors are reported in brackets. Panel A reports the results for India, while Panel B reports the results pooling the set of developing countries in Figure 2 together. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table B.3: Effect of Fertility on the Distribution of Human Capital in the Family - IV

# C Selected Survey Questions

This section report selected survey questions used in the survey experiment.

### Script for Beliefs

*We are interested in your opinion about how important it is for parents to devote resources to help their children acquire new skills. For this purpose, we will ask you to imagine an typical family that lives in a basi/neighbourhood like your own. The family has two children, Abhisekh and Biswajeet, and makes decisions about how much money to spend on educational resources that help their children acquire new skills and progress in their education (such as school fees, uniforms, books and school supplies, and private tutoring). We will show you different scenarios and ask you what you think the average monthly earnings of Abhisekh and Biswajeet will be at age 30 under each scenario. We will also ask you what grade you would expect Abhisekh and Biswajeet to reach in each scenario.*

*We know these questions are not easy to answer. Note that there is no right or wrong answer, we are just interested in what you personally think. Please try to consider each scenario carefully and tell us what you believe the outcome will be.*

Instruction for Interviewer: show VISUAL AID 0 to the respondent. Explain that the ruler represents children schooling ability. Worse children in school are at the bottom of the ruler while best children are at the top. Probe respondent understanding of the ruler by asking: *"Show me by pointing with your finger where the worse performing student in the school would be on this ruler?"*, and *"Show me by pointing with your finger where an average performing student in the school would be on this ruler?"*. If respondent shows understanding continue with the survey, otherwise continue explaining [the visual aid] until respondent understands.

Instruction for Interviewer: show VISUAL AID 1 to the respondent. Explain the scenario with the help of the visual aid. Explain that the arrows below the ruler indicate the positioning of Abhisekh and Biswajeet on this ruler.

*Abhisekh and Biswajeet are two healthy primary school aged children who attend the same school. At the beginning of the school year Abhisekh is among the top three students in his school and Biswajeet is among the bottom three students in his school.*

Instruction for Interviewer: while you go through the scenario show Abhisekh and Biswajeet position on the ruler by pointing with your finger.

Instruction for Interviewer: show VISUAL AID 1 to the respondent together with box A.

*A) If the parents spend 10 RUPEES every month on educational resources to help Abhisekh with his education, and they spend 10 RUPEES every month on educational resources to help Biswajeet with his education:*

- *How much do you think ... will earn on average per month at age 30?*
- *What grade would you expect ... to achieve?*

Instruction for Interviewer: show VISUAL AID 1 to the respondent together with box B.

*B) If the parents spend 10 RUPEES every month on educational resources to help Abhisekh with his education, and they spend 1000 RUPEES every month on educational resources to help Biswajeet with his education.*

- *How much do you think ... will earn on average per month at age 30?*
- *What grade would you expect ... to achieve?*

Instruction for Interviewer: show VISUAL AID 1 to the respondent together with box C.

*C) If the parents spend 1000 RUPEES every month on educational resources to help Abhisekh with his education, and they spend 10 RUPEES every month on educational resources to help Biswajeet with his education.*

- *How much do you think ... will earn on average per month at age 30?*
- *What grade would you expect ... to achieve?*

Instruction for Interviewer: show VISUAL AID 1 to the respondent together with box D.

*D) If the parents spend 1000 RUPEES every month on educational resources to help Abhisekh with his education, and they spend 1000 RUPEES every month on educational resources to help Biswajeet with his education.*

- *How much do you think ... will earn on average per month at age 30?*
- *What grade would you expect ... to achieve?*

## Script for Allocation Choices

*Now we will play a game with the goal of understanding how parents make decisions concerning their children, particularly how they make investment decisions in their education. We understand that these decisions are often very complicated and we are just interested in finding out more about what factors are important in these decisions. There are no right or wrong answers here and there is no intention to make any judgement.*

*We will present you another family who lives in a basi/neighbourhood like your own. This family has two children and decides how to invest some money on each of their children's education. The family asks for your advice on how to spend this money. We will tell you different stories and in each of these stories we will ask you to advice this family on how to invest in their children's education reflecting your choices.*

*The game has several rounds that correspond to different stories. In each round I will give you some beans that represent Rupees that the family has decided to spend on their children's education. Each story will be characterized by:*

1. *A total amount of Rupees to be spent. This is given by the total amount of beans.*
2. *An initial level of schooling ability of the two children.*

*After describing each story, I will ask you to advice the family on how to divide this money among their children (for example to pay for school fees, private tuition, or schooling materials). Please use* *the beans and place them in the boxes to reflect your choices. For example if you wish to assign all the resources to "Child 1" you should put all the beans in the box labelled "Child 1". Please notice that you have to place all the beans that I give you into the boxes. Let's practice with an example!*

Instruction for Interviewer: show VISUAL AID 4 to the respondent and hand 10 beans.
Trial 1: Probe respondent understanding by asking: *"Show me by placing the beans into the boxes how you would place the beans if you wished to spend all the rupees on Child 1."*
If respondent shows understanding continue, otherwise continue explaining until respondent understands.
Trial 2: Probe respondent understanding by asking: *"Show me by placing the beans into the boxes how you would place the beans if you wished to spend the same amount on both children."*
If respondent shows understanding continue, otherwise explain again until respondent understands.
Once you are confident that the respondent understands collect all the beans and move on.

*Please do not worry, there is no right or wrong answer and the intention is not to make any judgment. We understand that some of these questions might be hard, but please try to consider each scenario carefully. Before we start, do you have any question? Ok, let's start!*

*Imagine a typical family that lives in a village/neighbourhood like your own. The family has 2 primary school aged children, Pradeep and Sisir. At the beginning of the school year they decide how to spend some of their money on educational resources that will help their children to acquire new skills and progress in their education. The family asks for your advice on how to spend this money.*

*A) The family can spend 10 beans on their children's education. Pradeep and Sisir are both healthy children. At the beginning of the school year Pradeep is among the top three students in his school and Sisir is among the bottom three students in his school. I would like you to think about how this scenario and to place the beans into the boxes to reflect your choices.*

*B) The family can spend 10 beans on their children's education. Pradeep and Sisir are both healthy children. At the beginning of the school year Pradeep is among the top three students in his school and Sisir is an average student in his school.*