IFS

Institute for Fiscal Studies

Michele Giannola

Working paper 22/28

Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Lab-in-the-Field Experiment

Economic and Social Research Council

Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Lab-in-the-Field Experiment

Michele Giannola*

This version: July 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 lab-in-the-field experiment with poor parents in India. 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 outcomes, as well as study the role of household resources. I find that parents are driven by efficiency considerations rather than inequality concerns over children's final outcomes. Because they perceive investments and baseline ability to be complements in the production function, they invest more in higher-achieving children. Resources are important, as constrained parents select more unequal allocations. I then show that primitive parameters identified in the experiment are predictive of actual investment behaviour. The results indicate that families act as a reinforcing agent, magnifying ability-based educational inequalities between children.

*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 and Yale 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.

1 Introduction

Intra-household inequality is key for the measurement and understanding of poverty and inequality in a society (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 *between* children's educational outcomes, and studies the role played by parental investment to explain this inequality. Do parents invest more in higher-achieving children, exacerbating inequality in a society, or do they compensate 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 investment 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 wellbeing and reducing inequalities between children. 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 between children is plagued by the fact that measures of endowments, at birth or early in life, include both a genetic component *and* an endogenous 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. For example, a preference for a specific child would manifest itself in an higher endowment for the preferred child *and* a higher level of investment in that same child.[2]

Identifying the separate role of household preferences, beliefs and resources in the reduced-form demand equations for child human capital inputs is also challenging. The reason for this challenge is

[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)).

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

a twofold identification issue. First, as beliefs are usually not observed in standard survey data, input choices are consistent with many alternative specifications of preferences and beliefs (Manski (2004)). 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 resources to households might depend on how parents allocate those resources to children. For some specifications of parents' preferences, this policy could increase inequality between siblings (Barrera-Osorio, Bertrand, Linden, & Perez-Calle (2011)).

In this paper, I report the results of a lab-in-the-field experiment, purposely designed to understand parents' human capital investment in their children, that I conducted with poor parents in the urban slums of Cuttack, Odisha, India. I develop new theory-driven survey measures based on hypothetical scenarios that allow me to identify the causal effect of child human capital endowment on parents' investment. The strategically-designed survey instruments further allow me to separately identify the primitive parameters in the reduced-form demand functions: parental 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 then complement these strategically designed instruments with available behavioural data to validate my experimental strategy.

The 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 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 believe the future 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 complementarity or substitutability.

Having identified parental 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 investment decisions regarding their children's education, and asked to choose their preferred allocation of investment between two children with varying baseline ability. The experiment introduces exogenous variation in child endowment that I use to identify the reduced-form demand equations relating child ability to parents' choices. 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)), but upon which earlier work relies (e.g. Behrman, Pollak, & Taubman (1982)).[3]

The approach I use for identification has two other important advantages. First, it limits the role of parental preferences over other child attributes, such as gender (Barcellos, Carvalho, & Lleras-Muney (2014)), or birth order (Jayachandran & Pande (2017)), in driving parents' investment decisions. By using hypothetical scenarios, I can fix these child characteristics, as well as other features of the environment that might be unobserved by 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.

Several key results emerge from this study. First, I find that parents perceive baseline ability and educational investment to be relevant in the child 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 boosts earnings by 28 percent. Moreover, parents perceive the two inputs to be *complements*: They believe that investment is 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. This result speaks to a growing literature focusing on the role of subjective beliefs as a determinant of human capital investment decisions.[4] In particular, the finding suggests that parents' beliefs about the human capital production function matter to explain differences in investment between children *within* the same family, beside

[3] A recent exception is Berry, Dizon-Ross, & Jagnani (2020), which also uses a lab-in-the-field experiment to study parental preferences for investing in their children. To avoid making assumption on the perceived education production function, Berry, Dizon-Ross, & Jagnani (2020) shocks the *short-run* returns of investing in different children for identification, and studies how parents allocate lottery tickets for a one-hour tutoring. Differently from Berry, Dizon-Ross, & Jagnani (2020), this paper considers the role of parents' beliefs and resources in the human capital investment decision.

[4] See Attanasio & Kaufmann (2009); Jensen (2010); Arcidiacono, Hotz, & Kang (2012); Stinebrickner & Stinebrickner (2014); Wiswall & Zafar (2015); Boneva & Rauh (2017); Delavande & Zafar (2019).

their importance to explain inequalities *between* families (Boneva & Rauh (2018); Attanasio, Cunha, & Jervis (2019); List, Pernaudet, & Suskind (2021)).

Second, the experimental results reveal that parents reinforce baseline differences between children, and imply that, in this setting, parents have a low aversion to inequality over their children's human capital outcomes. Specifically, I show that when the difference in children's baseline ability increases, parents re-allocate towards the higher-achieving child. 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 by inequality concerns over final outcomes (Becker & Tomes (1976); Griliches (1979); Behrman, Pollak, & Taubman (1982); Behrman (1988); Pitt, Rosenzweig, & Hassan (1990)). Importantly, parental preferences are identified under far fewer assumptions than would be required from observational data.

Third, I document and important link between poverty and intra-household allocations. When parents have fewer resources, they make human capital investment decisions that are more unequal. This result suggests that poverty might prevent parents from adequately investing in *all* their children, leading them to select their *"strongest"* child, and leaving the more vulnerable children at 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, leading to a reduction in intra-household inequality. By showing that household resource have important implications for the allocation of human capital investments *between* children, this finding complements a literature investigating the role of resources in explaining the socio-economic gap in school enrolment and educational investments (Lochner & Monge-Naranjo (2012); Kaufmann (2014); Solis (2017)).

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)). One implicit assumption about this methodology is that elicited preferences are reflective of what respondents would do in the *real world*.[5] To assess the validity of this assumption in my setting, I exploit a unique feature of the survey, in which parents are asked detailed information on educational invest-

[5]Growing evidence points to the fact that the two approaches of using stated or actual choices yield similar preference estimates in a variety of contexts, especially when the hypothetical scenarios are realistic and relevant for the respondent (Mas & Pallais (2017); Wiswall & Zafar (2018)).

ments separately for each of their children in terms of both monetary and time investment.[6] I find a robust relation between elicited preferences and *actual* investment. In particular, respondents that are identified as less inequality averse invest more unequally in their children's education, favouring their higher-ability child. I also find that parents spend more productive time with this child. These findings suggest that these investment decisions might extend beyond educational expenditure to high-stakes investment choices that can have important long-term consequences for child wellbeing.

Finally, this paper contributes to a growing body of evidence pointing at the importance of considering intra-household inequality to understand differences across individuals in a society, and predict the effects of polices on individual welfare.[7] In particular, I simulate the effect of an intervention targeted at the household level (the level of targeting often used by interventions aimed at increasing parents' investment in child human capital) and 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 to the policy. This counterfactual experiment highlights that, to the extent that families are the ultimate decision makers, it is necessary to consider household behavioural responses to understand the impacts of policies on the wellbeing of individual household members.

The remainder of the paper is organized as follows. The following section presents two stylized facts that motivate this study. Section 3 presents the conceptual framework 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. Section 8 concludes.

2 Motivating Evidence

Two main 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

[6]This represents an improvement over standard household surveys, which collect information at the level of the household as a whole (e.g. educational expenditure for all children). Therefore, I can link important human capital inputs to each child, and study whether the behaviour of the parents in the experiment resembles how they invest in their own children.

[7]See for example 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). While this literature largely focuses on the understanding of 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.

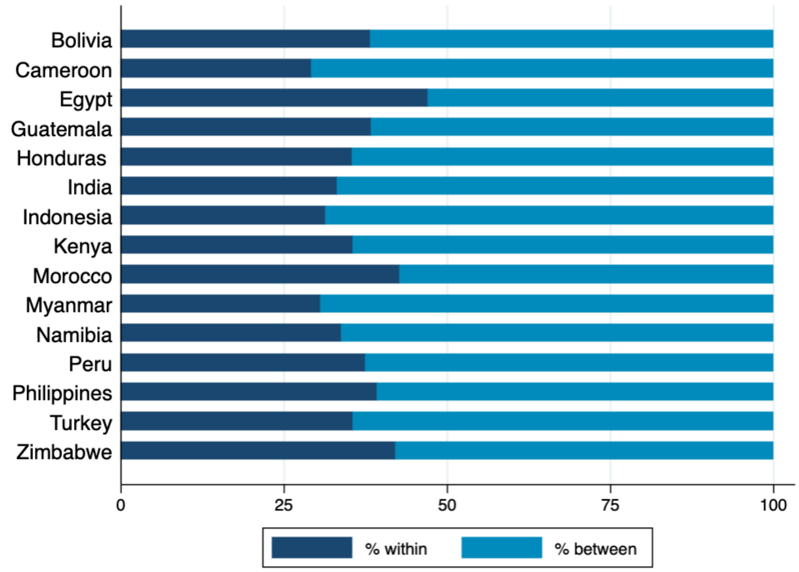


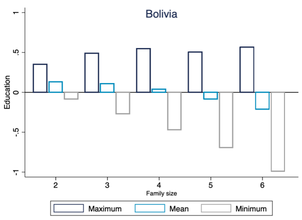
Figure 1: Contribution of Intra-household Inequality to Child Human Capital Inequality

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.

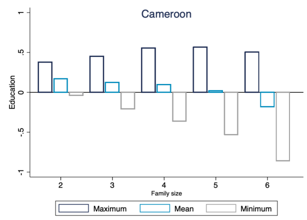
be attributed to within-household and between-households variation across a number of developing countries. To perform this decomposition, I use the Mean Log Deviation (MLD) measure of inequality (Ravallion (2015)), which can be exactly separated into a within-group component and a between-groups component (see Appendix A). The figure shows that, for a large set of developing countries, intra-household 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 E.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, the maximum, and the minimum of this distribution, that is the level of human capital of the *highest* and *lowest-achieving* children in the household, and the *average* level of human capital in the family. 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).

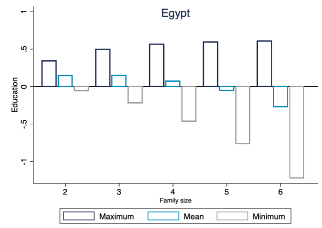
The figure reveals several patterns that are strikingly similar across countries. First, there is a negative relation between average child outcomes and family size. This is a relatively well documented



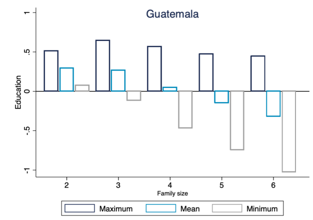
(A) Bolivia



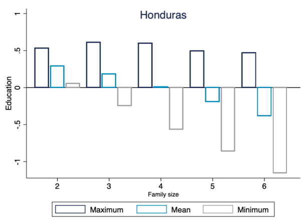
(B) Cameroon



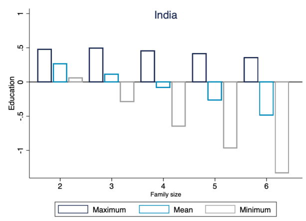
(C) Egypt



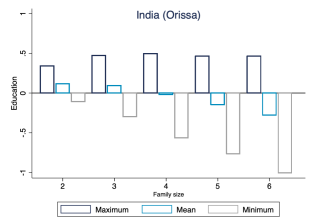
(D) Guatemala



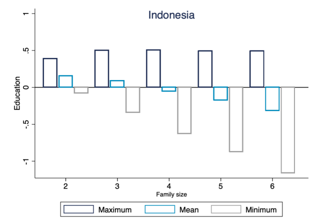
(E) Honduras



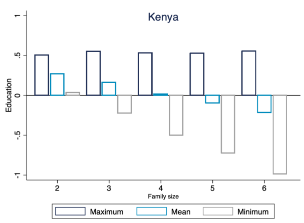
(F) India (whole)



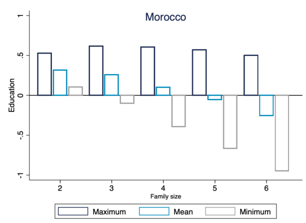
(G) Odisha (India)



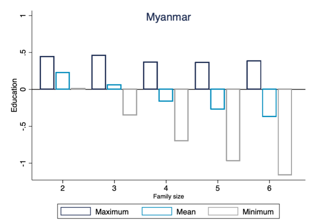
(H) Indonesia



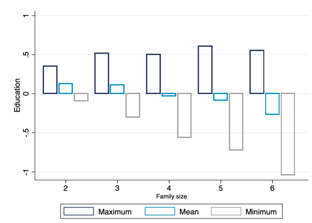
(I) Kenya



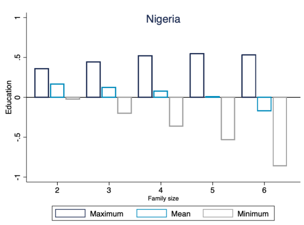
(J) Morocco



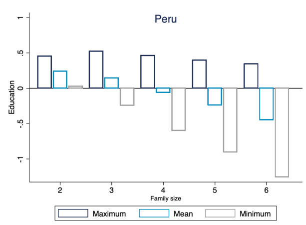
(K) Myanmar



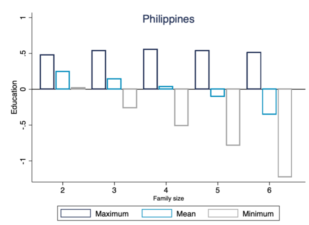
(L) Namibia



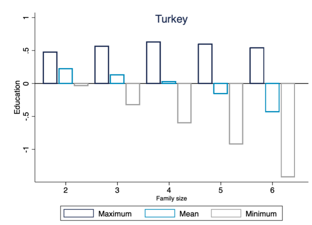
(M) Nigeria



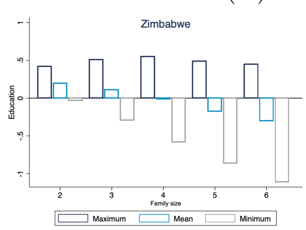
(N) Peru



(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 (light blue), the maximum (dark blue) and the minimum (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.

fact that dates back to Gary Becker's Quality-Quantity 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 Appendix B), including controlling for child gender and birth order, as well as considering the endogenous selections of families into different 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.[8]

The figure paints a more nuanced picture of the relation between child human capital and famiy size. While the Q-Q model implicitly assumes that parents invest similarly in *all* children, these patterns seem more 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: parents focus their investments on some children, and allocate the remaining resources to other children in the household. This investment strategy might be particularly detrimental for the human capital of children in larger families, because of less per-capita resources. Unequal investment between children 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 capial outcomes, this paper analyses empirically 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 simple framework to inform the design of the survey instruments used in the field, guide the empirical analysis, and interpret the findings.

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

3.1 Preferences and constraints

Parents derive utility from their children's human capital outcomes according to a Constant Elasticity of Substitution (CES) utility function that is 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 (e.g. her adult earnings or educational attainment), c_i is a child-specific preference that might depend on the child characteristics (e.g. a preference for sons over daughters), and ρ is the preference parameter that regulates parents' aversion to inequality in child outcomes. This functional form assumption is standard in the literature on intra-household allocation of resources (see, for example, 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. In particular, at one extreme when $\rho = 1$, the indifference curves are linear and there are no inequality concerns. In this case parents act 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.[9]

Parents choose educational investment in their children I_i to maximize their utility subject to two constraints. The first is a budget constaint. As this is a 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 stage budgeting process: in the first stage parents decide the amount of resources to spend on their children's education, and then decide how to distribute these resources between children.[10] The

[9]This specifications relies on several simplifying assumptions. First, I assume that the total number of children the family has is exogenous. Second, I do not consider the decision on how parents allocate resources between themselves and their children. Third, I consider an unitary model of the family where parents act as a single optimizing agent. I discuss some of these assumptions and how they can be relaxed in Appendix C. Also, as it is common in the literature, 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)). Future work should explore the possibility of ex-post equalization in this setting.

[10]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.

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.[11]

3.2 Subjective beliefs

Standard models of intra-household allocations of resources rely on strong assumptions about parents' knowledge of the human capital production function. In particular, these models often 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 practice. For instance, Cunha, Elo, & Culhane (2013), Boneva & Rauh (2018) and Attanasio, Cunha, & Jervis (2019) show that parents hold inaccurate beliefs about the productivity of different inputs entering the human capital production function.

To incorporate these information frictions into the model, I introduce the *perceived* human capital production function, which is specified as:

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

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 in the model. 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.[12] The maximization problem of the parents results in a policy function I_i^* describing the optimal level of

[11] Future work could extend this analysis to other dimensions of child endowments such as physical health. In practice, what matters to study parental behaviour is that child endowment affects the returns to investment, and is relevant to determine child human capital.

[12] I do not consider the issue of 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 this question. 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 their beliefs about the process of child development, and how these beliefs change over time should be the focus of future research.

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 issues

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 α and β are the returns to ability and investment. I assume that the perceived production function 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.[13]

Solving for the optimal level of investment in each child (see Appendix C), 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 optimal allocation of investments to parents' preferences (ρ, 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 from standard survey data, as these data usually only include information on investments and some proxies of children's endowments.[14] The corresponding

[13] 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 approximation 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 can also well approximate the perceive production function. Importantly, Cobb-Douglas production allows to derive closed form solutions for parents' investment and (as explained in Section 4) to point identify parents' inequality aversion, a key parameter in this literature.

[14] 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.

reduced-form equation, which one could estimate with observational data is:

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

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

1. The reduced-form coefficient γ is generally biased. Identification of the reduced-form parameter γ requires that differences in children's abilities are orthogonal to the error term ϵ 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).[15] For instance, a preference for a specific child would manifest itself in an higher endowment for the preferred child *and* a higher level of investment in that same child, resulting in a positive bias in the reduced-form coefficient γ. In general, a regression of parents' investment on children's ability does not identify the reduced-from coefficient γ.

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 choice data, as these are consistent with many alternative specifications of preferences and beliefs. Without imposing strong assumption on such 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 have some form of aversion to inequality in child outcomes, or that they believe that the returns to baseline ability are low: in both situations the corresponding reduced-form parameter would be close to zero.

As detailed in the next section, I overcome these identification challenges by implementing a lab-in-

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

the-field experiment with poor parents in Odisha, India. I design strategic survey questions (Ameriks, Briggs, Caplin, Lee, et al. (2020)) based on hypothetical scenarios, 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 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.

4 Lab-in-the-Field 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 lab-in-the-field experiment with parents of primary school children in urban Odisha, India. In the experiment, I use novel strategically-designed survey measures closely 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.[16] This strategy consists in presenting a series of hypothetical situations to the respondent and elicit information on individual expected outcomes.

The experimental procedures worked as follows. Surveyors presented respondents with a series of hypothetical stories (scenarios) about a representative family that lives in a neighbourhood like their own. The family has *two* children and makes decisions regarding their education. Guided by the theoretical framework, I focus on the role of perceived returns to child baseline ability (A_i), parental

[16]See Boneva & Rauh (2018), Attanasio, Boneva, & Rauh (2019) and Attanasio, Cunha, & Jervis (2019) for recent applications of this method.

investment (I_i), and on their perceived complementarity or substitutability. To identify these perceived returns, in each scenario I exogenously vary one of these characteristics and ask the respondent to report what they believe the future earnings of the child would be at age 30 (this corresponds to H_i in the theoretical framework).

The two children in the scenarios are described as attending the same primary school and identical in many respects, but differing in *one* important characteristic: while the first child - *Child H* - has an high baseline ability, the second child - *Child L* - has a low baseline ability. Specifically, to convey information about child 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, parents in India consider school performance as an important indicator of their child academic ability. This is also the first reliable and objective measure of child ability that parents have access to.

Scenarios then varied in 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 my experimental sample). Surveyors emphasized that these were *long-run* investments that would help the children acquire new skills and progress through their educational careers. Some scenarios described a *high* level of parental 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).

After presenting each scenario, surveyors asked respondents to report what they believed the likely outcome would be for each child, in terms of expected earnings at age 30. The respondent's answer was recorded, and the enumerator moved on to the next scenario. To help understanding, all scenarios were presented to the respondent with the help of a visual aid that sketched the main features and made salient to the respondent the differences across scenarios (Appendix Figure E.2 presents one of the visual aids used in the field, and Appendix F reports the exact wording of some relevant questions used in the survey). As a robustness check, parents were also asked to state what they believed the educational attainment of the child would be in each hypothetical scenario.

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 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). Presenting respondents with one hypothetical 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 might matter for child outcomes and that 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 that what matters is only the difference between the two children.

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 particularly important 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 child.

Finally, to understand whether average parental beliefs about the returns to child ability and investment significantly differ by child gender, respondents were randomized 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 perceived production function. Comparing responses across scenarios and between children I identify: (i) the perceived returns to investment, (ii) the perceived returns to baseline 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.

To characterise the perceived production function of 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 η_i are respondent fixed effects. The coefficients α_1 and α_2 identify the perceived returns to baseline ability and investment, while the coefficient α_3 identifies their perceived complementarity ($\alpha_3 > 0$) or substitutability ($\alpha_3 < 0$). Variants of this specification allow me to study whether *on average* perceived returns vary by child gender, by comparing respondents randomized in one of the two groups.

4.2 Second-stage experiment: Investment choices

Measurement. Having identified parents' beliefs, in the second round of the experiment I collect stated investment choices. As in the case of beliefs, parents are presented with a series of hypothetical scenarios. But in this stage of the experiment, instead of asking respondents to report what they believe the outcome of the child would be, they are 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 are presented with a representative family deciding how to distribute educational resources between their children. The resources being allocated are 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 and part of their educational budget). Similarly to the first-stage experiment, the survey script emphasized that these were *long-run* investments that could help the children acquire new skills and progress through their educational career.[17] Scenarios varied in terms of the following two characteristics: (i) children's baseline ability, and (ii) the total amount of resources that the family could allocate to their children's education.

Notice that while the equilibrium allocation rule in (8) does not allow to separately identify the primitive parameters in the demand function, it still provides some 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 introduce exogenous

[17] The exact wording of some relevant questions used in the survey is presented in Appendix F.

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 emphasized what different allocations would achieve in terms of total returns or cross-siblings 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.[18]

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. 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 fixed by design of the hypothetical scenarios.

The experimental design has two other important advantages. First, 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.[19] Second, the identification strategy is robust to parents having inaccurate beliefs about their children's baseline ability (Dizon-Ross (2019)), as these abilities are described to the respondent in each scenario.

To study whether parents' investment reinforce or compensate baseline differences between children,

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

[19] This is also important because standard household surveys usually 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.

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 β_1 pins down whether parents' investment is reinforcing ($\beta_1 > 0$) or compensating ($\beta_1 < 0$).

Because for many parents an important constraint to invest in their children might be the availability of material resources, I also test whether household resources matter to explain allocations. To do this, 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 β_3 identifies if reinforcement (compensation) is weaker when resources are higher (lower) ($\beta_3 < 0$).

Finally, as a previous work suggests that Indian boys and girls might be treated differently in terms of human capital investment, respondents were randomized 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. By comparing investment allocations between the two groups (holding fixed other child characteristics), I isolate 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 abilites, 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 of the hypothetical scenarios:

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

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 γ is identified from data on experimental allocations. Once these parameters are identified, parents' preferences can be recovered as:

$$\rho = \frac{1}{a} \times \left[\frac{1}{\gamma} + \frac{b}{a}\right]^{-1} \tag{12}$$

A consistent estimator for ρ can then be obtained by replacing the parameters in (12) with the corresponding OLS estimates from equations (9) and (10).[20]

Average preferences for child gender can be inferred by comparing the answers of respondents who saw two boys in the scenarios, with the answers of those who saw a girl and a boy.

4.4 Stated and Revealed Preferences

One question is whether preferences recovered from data on hypothetical choices relate to *real world* behaviour. To address this question, I also collect data on actual investments made by parents in the form of child specific inputs, and investigate whether respondents predicted to be less inequality averse in the experiment systematically make more unequal choices when it comes to distribute actual resources between their own children. Evidence in favour of this relation would add credibility to the research design, and to the use of hypothetical scenarios to identify key parameters of interest.

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% 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 for disadvantage children (see Andrew et al. (2019) for the first follow-up results).[21] 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% lived below the poverty

[20] Given consistency of the OLS estimator, using the continuous mapping theorem and Slutsky theorem one can prove consistency of the estimator for ρ. I obtain standard errors and confidence intervals for the preference parameter using bootstrap methods.

[21] Andrew et al. (2019) reports the short-run results of the psychosocial stimulation, and find positive effect on child cognitive development 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†	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†	6.662	9.555
C. Household members' characterstics		
Household member age	26.129	18.538
Household member is male	0.481	0.500
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. † indicates expenditure in thousands of INR. Educational expenditures include school tuition, money spent on purchasing textbooks and stationery, and hiring private tutors. The exchange rate was 71.43 INR : 1 USD at the time of the study.

Table 1: Summary Statistics

line. Households were then randomized 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. Appendix Tables D.1 and D.2 show that there are no effects of treatment allocation on parents' beliefs or preferences. These results from the first follow-up also showed that there were no improvements in maternal knowledge of child development (Andrew et al. (2019)). I thus ignore the treatment allocation and report results pooling the treatment and control groups together.

In 2019, we aimed at re-interviewing all households in the original sample.[22] Survey respondents were for the most part children's female primary caregivers, who were usually their mothers. The lab-in-the-field experiment took place in respondents' homes, during the caregivers' endline survey and,

[22]To increase the sample size (for the purpose of this 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).

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, and time investments by the parents such as time spent helping the child with homework), and child time use.

Table 1 reports the summary statistics for 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 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 is potentially very 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' 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 D.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 around 80 percent (columns 1 to 4). At the sample mean of expected earnings (30,120 INR) this corresponds to an increase of roughly INR 24,000. I discuss the magnitude of this coefficient below. Interestingly, when I control for expected years of schooling (as reported by the respondents) the coefficient on child baseline ability decreases by almost 50 percent (column 5). This is because, as show in Appendix Table D.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).

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 2 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 that the returns to their investment are 12.6 percent higher for a high-ability child than a 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.[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

[23]Results for educational attainment (reported in Appendix Table D.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.

	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
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

the coefficient on child gender in Table 2, and show in Appendix Table D.4), they do not perceive the returns to ability or investment to substantially differ between girls and boys (see Appendix Table D.4). 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 might 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.[24,25]

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 Attanasio, Boneva, & Rauh (2019) and construct an individual-specific measure of perceived returns. For example, I compute

[24]Essentially, this can be interpreted as a shift in the total factor productivity in the production function, while leaving unchanged the other technology parameters. Similar results using educational attainment as the outcome variable are reported in Appendix Table D.5.

[25]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 age 8. They also find that investment and baseline development are complements.

individual perceived returns to investment as the difference between a respondent's 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 results in Appendix Figure E.3. Panel A displays the empirical cumulative distributions of individual perceived returns to investments. The figure reveals a substantial variation in perceived returns across respondents: the $10th$ percentile is 0 and the $90th$ percentile is 0.48.[26] By comparing the expected earnings of the high and low ability child, while holding investment fixed, I also compute the individual perceived returns to child baseline ability. The distribution of these perceived returns is shown in panel B of Appendix Figure E.3, and also shows substantial heterogeneity: the $10th$ percentile is 0.33 and the $90th$ percentile is 1.19.[27]

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 resources allocated to the higher-ability child, which at the sample mean corresponds to a 14 percent increase in resources devoted to this child. The positive coefficient implies that parents' investment is *reinforcing*.

Table 3 also shows that household resources play an important role to explain 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*

[26] Figure E.3 also shows that, consistently with the findings form Table 2, the distribution of perceived returns to investment for the the higher-ability child first is shifted to the right compared to the distribution of perceived returns for the lower-ability child.

[27] I correlate perceived returns with observable characteristics in Figure E.5 and find that the education level of the primary caregiver predicts higher perceived returns to investment.

	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
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 randomized in seeing two boys and zero if the respondent was randomized in seeing one boy and one girl. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table 3: Intra-household Allocation of Resources

time of crisis" (Duflo (2005)). Therefore, it seems plausible that relaxing resource constraints could contribute to close investments gaps between children, potentially resulting in lower intra-household inequality in outcomes.

Interestingly, Table 3 shows no evidence that investment choices depend on the gender of the child.[28] 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 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.[29]

[28] I also tried estimating equations (10) and (11) separately for the two different groups defined based on the gender of the two children, and found very similar results.

[29] 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

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 shown in Table 2). One potential reason that could explain this result might be 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 for investing in a girl's education is a substantial perceived marriage market return to schooling (Adams-Prassl & Andrew (2020); Ashraf, Bau, Nunn, & Voena (2020)).

Parents' preferences. As discussed earlier, using the experimental allocations to regress parents' investment on child ability identifies the reduced-form parameter γ. 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 ρ that reconciles choices with parents' beliefs is positive and statistically significant at the 99% 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, this parameter is very close to the parameter estimate reported in Behrman (1988) for India (0.47). The estimated ρ is also statistically different from 1 at the 99% confidence level. This coefficient implies that parents weight relatively more efficiency than inequality-concerns, but that they are not pure returns-maximizers when investing in their children, as would implied by a value of ρ equal to 1.

Similarly to the case of beliefs, I also study heterogeneity in preferences. I plot the empirical cumulative distribution function of individual preferences in Appendix Figure E.4. Interestingly, for all families in the sample ρ is positive. However, some families are significantly less inequality averse than others (i.e. they have an higher value of ρ). I use this heterogeneity to classify families as *low* and *high* ρ *types* by splitting the sample at the median value of the distribution of ρ_i.[30,31]

be undernourished. Girls are also more likely than boys to be fully vaccinated (55% of girls, compared with 49% of boys) (IIPS (2001); IIPS (2008); Padhi (2001)).

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

[31] I correlate parental preferences with observable characteristics in Figure E.6 and find that families with more children are less inequality averse, while households attached to the sewage system are more inequality averse.

6.3 Stated and Revealed Preferences

The final question that I address is the relevance of the results outside the experiment and, in particular, whether elicited preferences reflect *real world* behaviour. To answer this question, I exploit a unique 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 collect information at the level of the household as a whole (e.g. educational expenditure for all children), or only on health inputs (e.g. breastfeeding or vaccinations which are common in large scale household surveys). I therefore use this rich information on current investments made by the respondents in their children, and relate these investments to child ability. 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.*"

Notice that what this questions captures is 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.

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 including age and gender. 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 INR 290 in favour of the higher-ability child. At the sample mean, this corresponds to 3.8 percent of total yearly educational expenditure. Appendix Table D.6 breaks down this results by different 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 INR 234 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 estimated ρ_i is above or below the

	Educational expenditure†			Child work			Private school			Time index		
	All	ρ_L	ρ_H	All	ρ_L	ρ_H	All	ρ_L	ρ_H	All	ρ_L	ρ_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 (ρ_L means higher inequality aversion). These two groups are defined based on whether the estimated ρ 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

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 INR 509. 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 IRN 240 (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 to identify primitive parameters of interest.

The remainder of the table shows results for additional investment measures. Columns 4 to 9 of Table 4 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% 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.

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 D.7 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). Interestingly, 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.

	U	I_H	I_L	H_H	H_L
	(1)	(2)	(3)	(4)	(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 ρ is set to its estimated value of 0.449 from section 6.2, while in Panels B and C report the results setting ρ 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

These additional results suggest that parents' investment decisions might expand beyond educational expenditure, with potentially important long-term effects for children's wellbeing.

7 Policy Experiment and Welfare

In this section, I explore the implications of my results to understand the effects of a policy designed to improve child outcomes and wellbeing. In particular, I consider the effects of an intervention that affects parents' information about the returns to investments in children, while holding fixed the educational budget. Such interventions have become increasingly popular in recent years both in developed and in developing countries, and are often targeted at disadvantaged families (see the discussion in Section 6.3 of Duncan, Kalil, Mogstad, & Rege (2022), and the references therein). Although these programs are usually targeted at the household level, they often have one "target child'. Consequently, evaluations of such interventions usually collect data only on this "target child". Here, I consider the effects of the policy on parents' investment in *each* child, on children's final outcomes, as well as on household *and* child welfare. I assume that child welfare corresponds to her level of human capital, while parents' welfare corresponds to their utility.

I assume that the family has only two children, and there are no child-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 $\theta_H > \theta_L$. I consider the effect of a policy that increases the *perceived* productivity of investment a, aligning it with its *true* productivity α, which is assumed to be larger so that $a < \alpha$. In the simulations, I keep fixed all other model parameters; in particular, I set the parents' preference parameter ρ 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 goes up after the policy is implemented (column 1). The magnitude of the effect of the policy on parents' utility is comparable to a 1% increase in the total educational budget y.[32] At the same time, because the intervention increases the perceived returns to investment, parents adjust their behavior investing more in their higher-achieving child and less in their lower-achieving child (columns 2 and 3). As a result, the policy improves the welfare of the higher-ability child, but decreases the welfare of the lower-ability child (columns 4 and 5). As shown in Panels B and C of Table 5, these results are robust to using different values of parents' inequality aversion. In these two panels, I report the same set of results, setting ρ equal to the lower and upper bounds of the 95% bootstrap confidence interval from Section 6.2.[33]

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 families are the ultimate decision makers, it is necessary to understand how they allocate resources to individual children and what determines their behavior 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 programs targeting the home environment, but might be important to gain a better

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

[33]More generally, re-optimization from the parents following the policy will occur for all values of ρ that are different from zero.

understanding of their (lack of) impacts.

8 Conclusions

This paper studies the role of parents' human capital investments as a determinant of intra-household inequality in child outcomes. I fist 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 lab-in-the-field experiment, motivated by a simple model of parents' human capital investments. 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 504 poor parents with children in the slums of Cuttack, 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 their investments. Second, I find that parents have a low aversion to inequality over their children's outcomes, so that 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 per-capita 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. Taken together, my results indicate that in this setting families act as a reinforcing agent, magnifying initial inequalities between children.

These findings have important implications for policy. First, the fact that parents respond to initial 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, thus magnifying total returns. My findings also point to the role that household resources have to explain human capital outcomes, and in particular they highlight a link between poverty and intra-household allocations. 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.

The results also suggest that the effects of any interventions aimed at improving children's outcomes will crucially depend on parental behavioural responses, which are mediated by their preferences and beliefs. To the extent that families are the ultimate decision makers, it is necessary to understand how parents allocate resources to individual children and what determines their behavior in order to predict the effects of policies and understand their impacts on *individual* children's wellbeing. 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. 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.

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.

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, Boneva, T., & Rauh, C. (2019). *Parental beliefs about returns to different types of investments in school children* (Tech. Rep.). National Bureau of Economic Research.

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., 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., 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.

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.

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., & 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.

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.

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

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.

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.

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.

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. (2021). *It all starts with beliefs: Addressing the roots of educational inequities by shifting parental beliefs* (Tech. Rep.). National Bureau of Economic Research.

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.

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.

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*.

Appendices

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:

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

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:

$$MLD = \sum_{j}\frac{N_j}{N}MLD_j + \sum_{j}\frac{N_j}{N}ln\frac{\bar{y}}{\bar{y_j}} \tag{A.2}$$

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

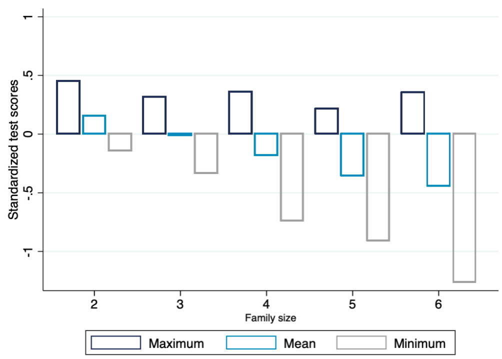
This section provides several robustness checks for the relation between fertility and the distribution of human capital in the family shown in Figure 2.

- 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.

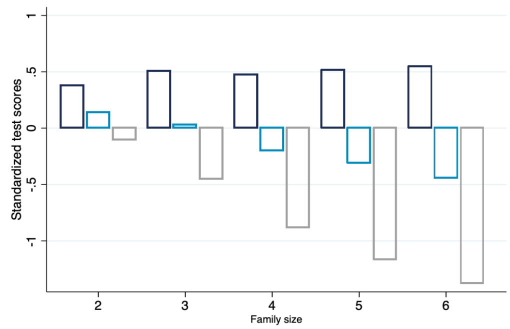
- Table B.1 report the regression results using age standardized test scores as measure of quality. In the table, I report the results of separate regression for the mean (columns 1 to 4), the maximum (columns 5 to 8) and the minimum (column 9 to 12). Columns 1, 5 and 9 include a linear indicator for 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.

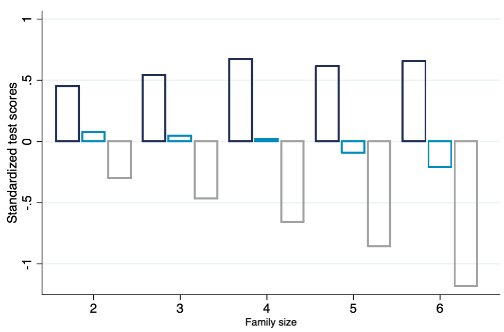
- Table B.3 reports the IV results using years of schooling as measure of human capital. Family size is instrumented using twin birth as an instrument for total family size. In the table, I report separate regressions for the mean, the maximum and the minimum. Panel A reports the results for India, while 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.



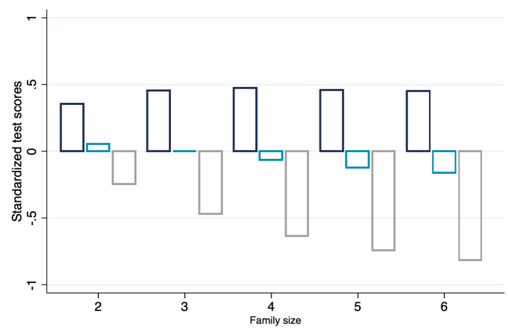
(A) India



(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 Model Appendix

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:

$$X_i^* = y \frac{c_i^{\frac{1}{1-b\rho}} \times \theta_i^{\frac{a\rho}{1-b\rho}}}{\sum_{j=1}^{n} c_j^{\frac{1}{1-b\rho}} \times \theta_j^{\frac{a\rho}{1-b\rho}}} \tag{C.3}$$

Computing the ratio of X_i^* to X_j^* and taking the log we get equation (7).

Human capital investments and family size

Holding fixed child specific preferences, for sufficiently high values of γ the highest ability child receives roughly the same share of household resources independently of family size. To see this define the highest ability child in the family as θ_{max}. For any family size n, the model implies that educational investments in the highest ability child are:

$$X_{max}^*(n) = y \frac{\theta_{max}^{\frac{a\rho}{1-b\rho}}}{\underbrace{\theta_{min}^{\frac{a\rho}{1-b\rho}} + ... + \theta_{max}^{\frac{a\rho}{1-b\rho}}}_{\text{n terms}}} = y \frac{\theta_{max}^{\gamma}}{\theta_{min}^{\gamma} + ... + \theta_{max}^{\gamma}} \tag{C.4}$$

As n increases, competition over household resources increases. This can be seen from the increase in the number of terms on the denominator of expression (C.4). Diving both the numerator and denominator by θ_{max}^{γ}, and taking the limit as $\gamma \to \infty$:

$$\lim_{\gamma \to \infty} X_{max}^*(n) = \lim_{\gamma \to \infty} y \frac{1}{\left(\frac{\theta_{min}}{\theta_{max}}\right)^{\gamma} + ... + 1} = y \tag{C.5}$$

This is because the first $n-1$ terms in the denominator are smaller than one. The result holds for all values of n so that $X_{max}^*(n) \to X_{max}^*(n+1)$.

Discussion of the model

This section discusses model assumptions and extensions.

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)).

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). Fertility choices depend on parental preferences for intra-household inequality (the parameter ρ). 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 θ_i.[34] Importantly, the optimal allocation rule is not affected by the fertility decision. 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.

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 high achieving children are not affected by variations in family size, the human capital of low achieving children sharply declines as family size increases. This heterogeneous effect of family size on child outcomes could potentially

[34] Using data from the Indian National Family and Health survey, I test and find empirical support for this model's prediction (results available upon request). Interestingly, 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)).

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 (**?**; 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 per-capita resources, combined with parental investment decisions.

D Appendix Tables

	Perceived earnings at age 30 (log)	
	(1)	(2)
High ability	0.897***	0.837***
	(0.046)	(0.046)
High Investment	0.234***	0.175***
	(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***
		(0.027)
High ability × High investment × Treatment		-0.038
		(0.043)
Boy	0.154***	0.154***
	(0.056)	(0.056)
Mean outcome	29676	29676
R^2	0.364	0.365
Observations	2480	2480

Notes: This table presents analogous coefficients and standard errors to those presented in Table 2 but with all the main regressors interacted with RCT treatment status. Because treatment status is allocated at the respondent level, regressions do not control for family fixed effects. The relevant comparison for column 1 is column 2 from Table 2 , and for column 2 is column 3 of Table 2. The number of observations is smaller because not all participants to the lab-in-the-field 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 D.1: Effect of RCT Treatment Status on Perceived Production Function

	Share of resources to child H	
	(1)	(2)
Difference in ability	0.077***	0.109***
	(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***
		(0.011)
Difference in ability × High resources		-0.064***
		(0.014)
High resources × Treatment		0.005
		(0.016)
Difference in ability × High resources × Treatment		0.032
		(0.021)
Boy	0.006	0.006
	(0.011)	(0.011)
R^2	0.076	0.087
Observations	1240	1240

Notes: This table presents analogous coefficients and standard errors to those presented in Table 3 but with all the main regressors interacted with RCT treatment status. Because treatment status is allocated at the respondent level, regressions do not control for family fixed effects. The relevant comparison for column 1 is column 1 from Table 3, and for column 2 is column 3 of Table 3 . The number of observations is smaller because not all participants to the lab-in-the-field 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 D.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 D.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 D.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 D.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 D.5: Perceived Production Function by Gender (Educational Attainment)

	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 D.6: Actual Parental Investments: Expenditure Category

	Encourage to read	Conversations	Outings	Play	Homework	Discuss shool
	(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 D.7: Actual Parental Investments: Time

E Appendix Figures

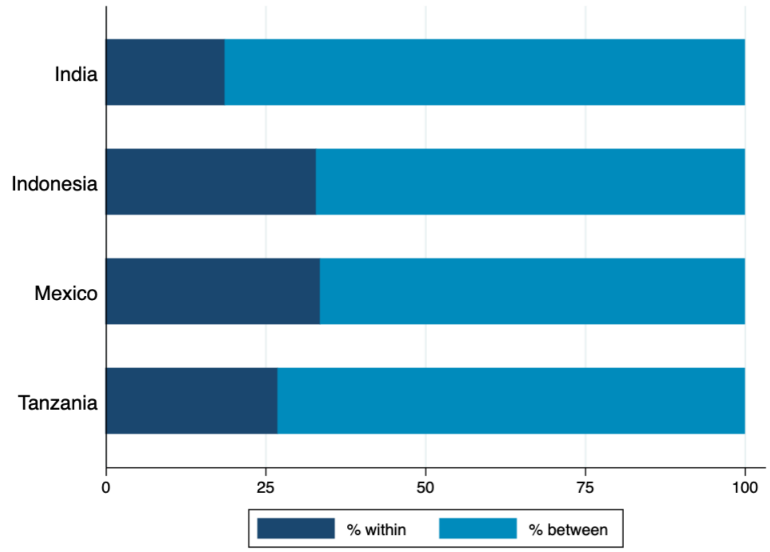


Figure E.1: Inequality in Child Human Capital (Test Scores)

Notes: This figure plots the within-household and between-households component of the Mean Log Deviation (MLD) measure of inequality. The outcome variable is age 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.

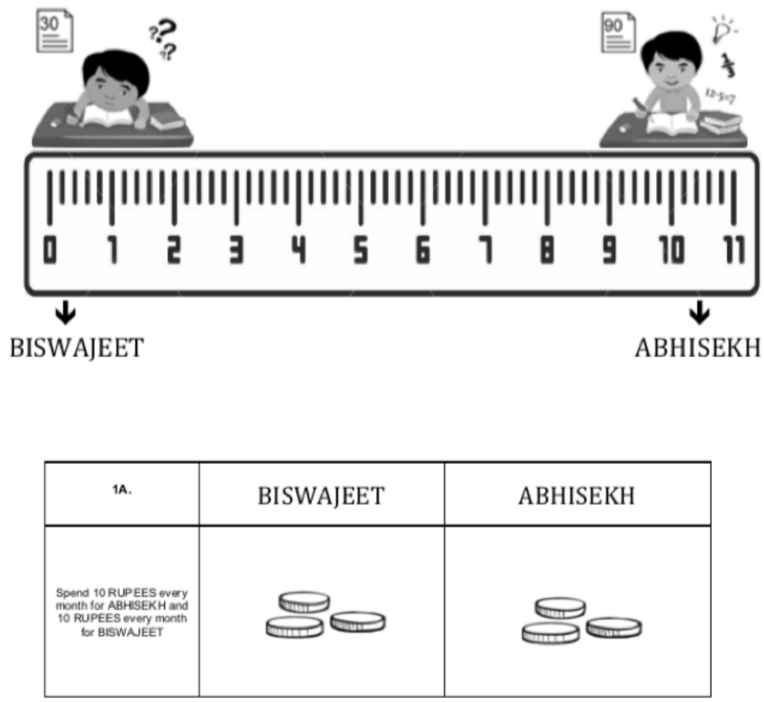
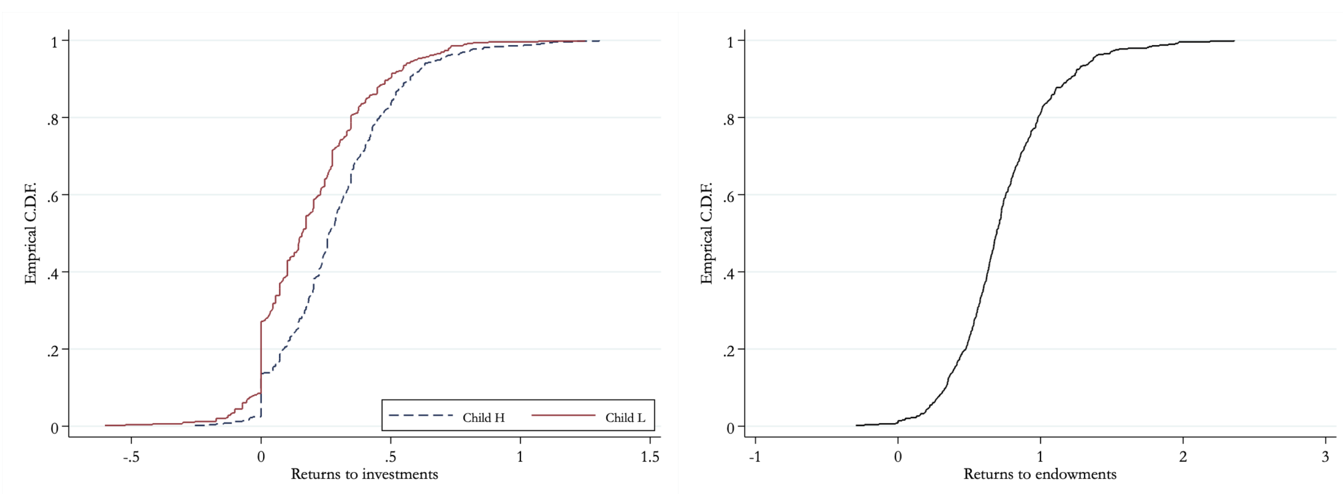


Figure E.2: Visual Aid

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 investments

(B) Returns to child ability

Figure E.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 investment, while panel B shows the CDF for the perceived return to child ability. Panel A 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.

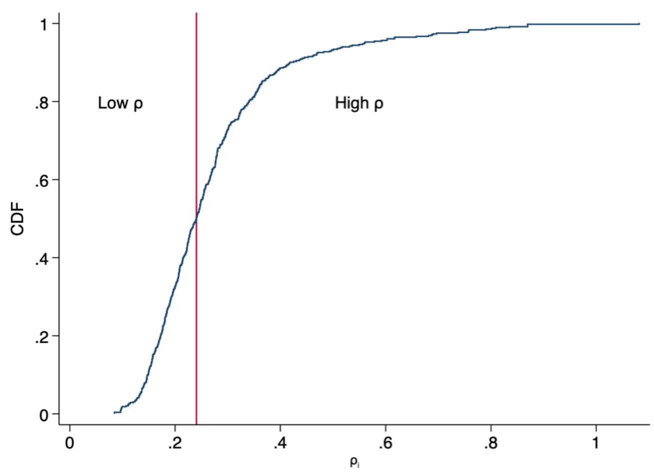
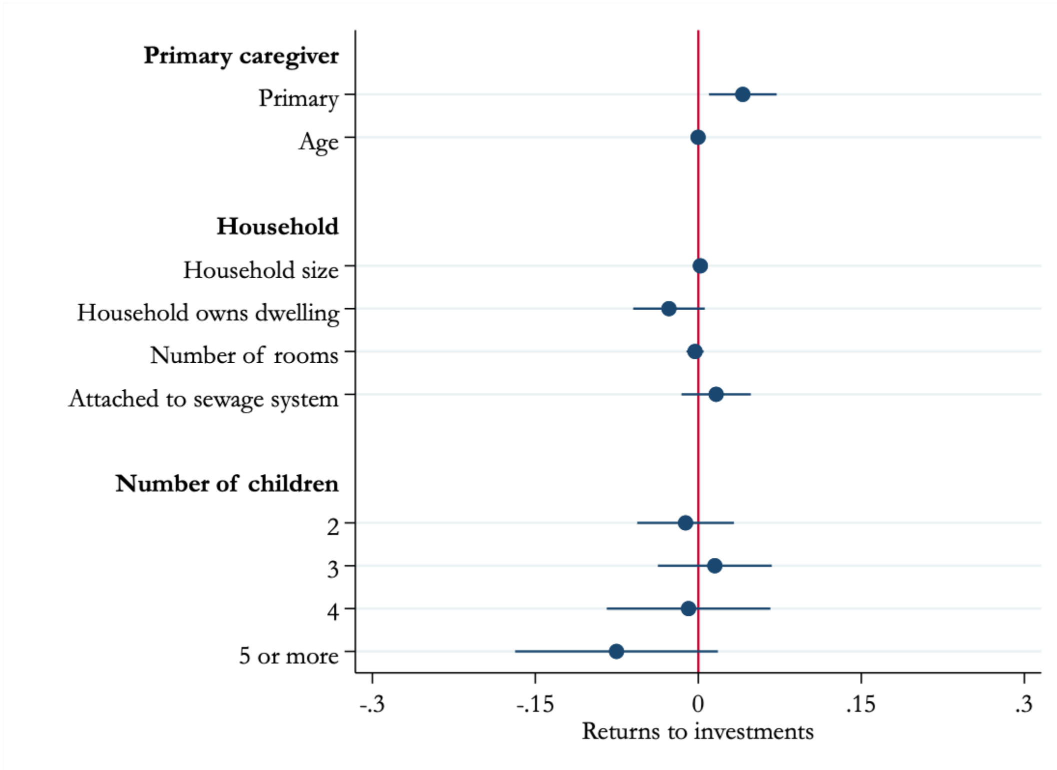
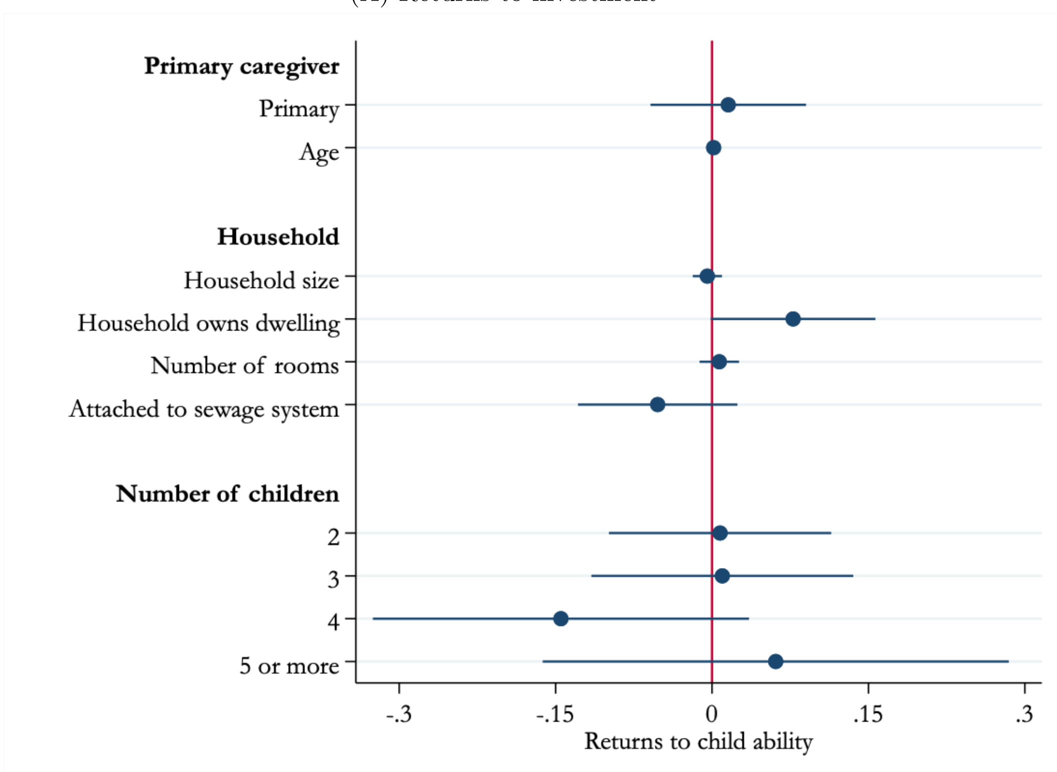


Figure E.4: Heterogeneity in Preferences

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



(A) Returns to investment



(B) Returns to child ability

Figure E.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) on observable characteristics.

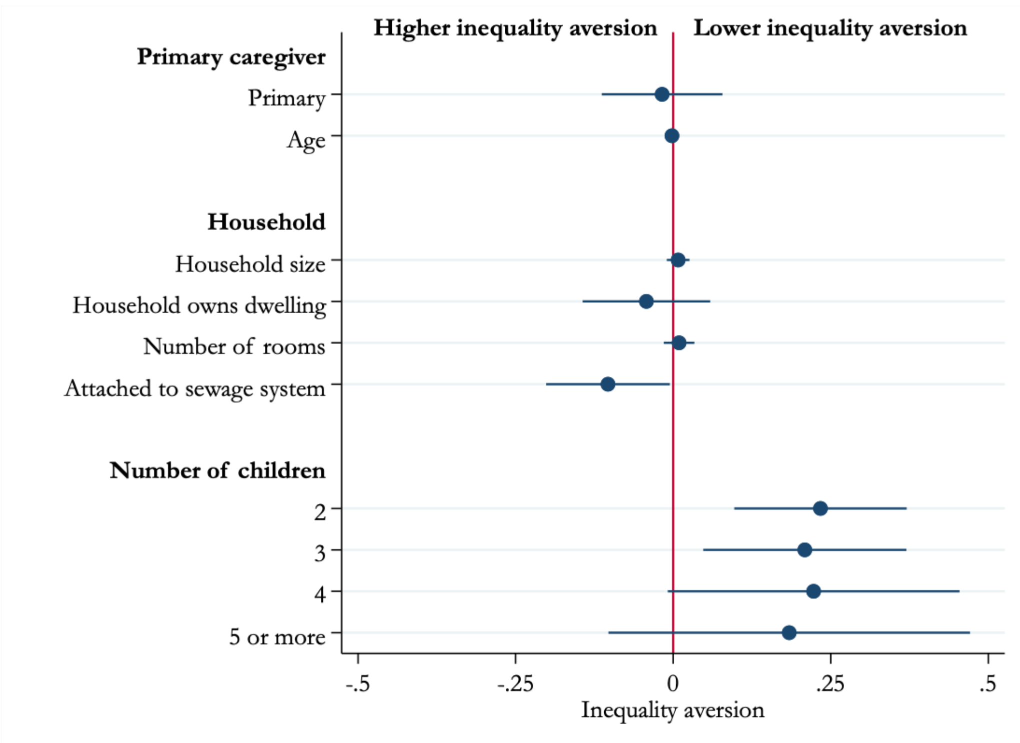


Figure E.6: Correlations between parental preferences and observable characteristics

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

F Selected Survey Questions

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. 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 (e.g. to pay for school fees, private tuition, schooling materials, etc.). 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 responder 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 responder shows understanding continue, otherwise explaining 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.