ESSAYS ON THE ECONOMICS OF HUMAN CAPITAL

Michele Giannola

March 2021

Submitted to the Department of Economics in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

UNIVERSITY COLLEGE LONDON

Declaration of Authorship

I, Michele Giannola confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

This thesis studies the process of human capital formation in the childhood years, focusing on the role that families, institutions and public interventions can play in this process. One chapter analyses the details of the process of child development, focusing on the evolution of and interaction between health, cognitive and socio-emotional skills. We find that skills are particularly malleable and parental investments highly productive in the early years. Investments depend strongly on household resources, and this is consistent with wealth gradients in child development. We also show that early health has long-term impacts on cognitive development. A second chapter combines a model of household behaviour with a lab-in-the-field experiment to study intra-household inequality in child human capital outcomes. The model highlights the key role that educational investments made by parents in their children can have to explain inequality, and how these investments depend on parental preferences, beliefs and constraints. To mitigate the identification problem posed by observational data, I collect original data on subjective expectations and stated choices through a lab-in-the-field experiment. I find that parents perceive child ability and investments to be complements. As parents have a low aversion to inequality in child outcomes, they reinforce initial differences across children. Household constraints are also important to explain choices. The last chapter of the thesis studies peer effects in academic achievement. Using the randomized evaluation of a remedying education intervention targeting low-achieving students within a class, we show that the test scores of their higher achieving classmates increased compared to similar children in control schools. We interpret this effect through the lens of a linear-in-means model of peer effects. The findings suggest that policies aimed at improving the bottom of the achievement distribution can generate social-multiplier effects benefiting all.

Impact Statement

This thesis studies the process of human capital formation in the childhood years, focusing particularly on the role that families, institutions and public interventions play in this process. The relevance of this topic has been recently highlighted by the fact that early childhood development is now explicitly recognized as a target (goal 4.2) under the United Nations sustainable development goals for 2030 (United Nations (2016)). The results presented here are therefore relevant for both the academic and policy debates.

Across developing countries, more than 200 million children aged less than five years are failing to reach their developmental potential because they suffer from the negative consequences of poverty, nutritional and health deficiencies, and inadequate learning opportunities (Grantham-McGregor et al. (2007)). Understanding the features of the process of human capital formation in the childhood years, such as its dynamic properties and the persistence of different inputs, is therefore key for the design of policies aimed at remedying early disadvantage. The first chapter of this thesis sheds light on the details of the process of human capital formation, providing one of the most detailed characterizations of this process. The results presented are important for the dynamic targeting of interventions, and the identification of *windows of opportunities* for child development. In particular, they highlight the potential that early investments have to improve the outcomes of disadvantaged children. The findings also demonstrate the existence of key complementarities between health and subsequent cognitive development, suggesting that preventing ill health from an early stage is likely to be an important element in the design of interventions aimed at improving children's outcomes in the long run.

The second chapter provides new evidence on the determinants of parental investments in their children's education, and on the implications that these choices might have for human capital inequalities in developing countries. The empirical analysis demonstrates that families respond to early levels of child development by investing more in higher ability children, thus amplifying initial inequalities. This has important implications for the understanding and measurement of inequality across individuals in a society. By showing that parental investments respond to child development, the results also highlight how the effect of public interventions crucially depends on parental endogenous responses, which are mediated by their preferences, beliefs and constraints. Incorporating these responses is likely to have an important impact for the design of effective interventions.

The final chapter provides the first successful example of how the existence of peer effects in education can be exploited in the design of policies aimed at improving students' academic performance. Our findings suggest that remedying education interventions aimed at improving the bottom of the achievement distribution can generate social-multiplier effects, so that it is possible to substantially improve the quality of education for all with relatively cheap and easy-to-scale interventions targeted to the weakest. The results provide a strong rationale that underscores why society should care about improving the educational outcomes of low-achieving students, and are likely to inform the policy debate concerned with the allocation of public funds to education.

Acknowledgements

I would like to express my sincere gratitude to my advisors Orazio Attanasio and Imran Rasul, whose guidance and support was essential throughout the PhD. Orazio has been the perfect primary advisor. He has always believed in my research ideas, and has encouraged me to push my own limits in order to pursue them. It has been an absolute pleasure to learn from and interact with him over these years. Imran has been the ideal secondary advisor, who constantly challenged me to think more deeply about any given problem. I am grateful for the countless productive interactions, which have significantly benefited the quality of my research.

I would like to thank the amazing team of researchers at the Centre for the Evaluation of Development Policies at the Institute for Fiscal Studies for involving me in many fascinating projects, and helping me collect the data that made part of this PhD thesis possible. In particular, Lina Cardona-Sosa provided excellent field assistance.

Over these years, I have also enjoyed very much the stimulating discussions with Samuel Berlinski, Raquel Bernal and Matias Busso.

I am grateful to the Economic and Social Research Council for the financial support received during my studies, and for allowing me to expand my research horizons by visiting the World Bank and the Inter-American Development Bank. Those visits have been instrumental to the development of this thesis, and essential for my training as an economist.

Many thanks to my friends Alessandra, Carlo, Divya, Federica, Julen, Manuela, Mimosa and Silvia. Sharing with you these years in London has been an absolute pleasure, and made this whole adventure a great experience.

A special thank you goes to my parents, Adriano and Lilia, and to my sister, Giulia, who have supported me as always throughout these years. Without them this accomplishment would not have been possible.

Lastly, I would like to thank Giulia, for her optimism, enthusiasm and positive attitude which helped me every step along the way. Her unconditional love and support have been essential throughout this journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Across developing countries, more than 200 million children aged less than five years are failing to reach their developmental potential at great cost to individuals and their countries. Poverty, poor health and nutrition, and inadequate care are the main risk factors (Grantham-McGregor et al. (2007)). Because the early childhood experiences are key for human capital development and well-being over the life-cycle, understanding what role families, institutions and early interventions play in this process is of paramount importance to tackle inequality in opportunities. At the same time, a better understanding of the features of the process of human capital formation is key to inform the design of appropriate policies to remedy early disadvantage. This thesis consists of three essays that examine the determinants of human capital outcomes in the childhood years, and what role public polices can have to improve children's outcomes, particularly for disadvantaged populations in developing countries.

The first essay "Child Development in the Early Years: Parental Investment and the Changing Dynamics of Different Dimensions", co-authored with Orazio Attanasio, Raquel Bernal and Milagros Nores, uses the data collected around the evaluation of a nursery program for disadvantaged children in Colombia to estimate the details of the process of human capital formation in the early years. We model child health, cognitive and socio-emotional skills, and study how these evolve over time and interact with each other as well as with other inputs – in particular parental investments – to determine the level of development at following ages.

We explicitly acknowledge the possibility that our measures of human capital and investments are contaminated by measurement error, and estimate dynamic latent factors models. We pay particular attention to the issues of scaling and normalization of the latent factors which are especially relevant when interest lies in understanding the dynamic properties of the processes (Agostinelli & Wiswall (2016a)). To this aim we develop a new procedure that allows comparison of the latent factors over time when no age-invariant measure is available over the whole period under study.

We find that the dynamics of the process can be richer than usually assumed. This has important implications for the degree of persistence of different inputs in time. In particular, we show that, compared to older ages, childrens skills are particularly malleable and parental investments highly productive in the early years. These results are important for the dynamic targeting of investment and interventions, and the identification of *windows of opportunities*. For instance, the results for cognitive skills and health suggest that investments taking place early in the life cycle are significantly more productive that later investments. At the same time, our results imply that even interventions with large positive effects in the short run can display some fadeout over time if not sustained and followed-up by later investments. We also find that parental investments depend strongly on household resources. This latter result is consistent with and can explain the existence of wealth gaps in child development. Finally, we show that early health has long-term impacts on cognitive development, suggesting that preventing ill health from an early stage is likely to be an important element in designing interventions aimed at improving child cognitive outcomes in the long run.

In the second essay "Parental Investments and Intra-household Inequality in Child Development: Theory, Measurement, and Evidence from a Lab-in-the-Field Experiment" I study how parents in developing countries make investment decisions in their children's education, and the implications that these investment choices have for intra-household inequality in child human capital outcomes. To this aim, I combine a theoretical model of household behaviour with experimental data collected through a lab-in-the-field experiment.

I use the theoretical framework to motivate my empirical analysis, guide the design of the survey strategy and measurement tool used in the field, and interpret the empirical findings. The model highlights the role that parental investments have to explain inequality in child outcomes, and how these investments depend on parental preferences for intra-household inequality, beliefs about the process of child development, and resource constraints. The model therefore shows the challenge that the use of observational data poses to the identification of key parameters and mechanisms of interest. To mitigate the identification problem posed by observational data, I design a novel measurement strategy based on hypothetical scenarios that allows me to elicit direct measures of parental beliefs, identify preferences for intra-household inequality, and study the role that resource constraints have to determine choices. I embed this in a lab-in-the- field experiment with parents of primary school children in India.

I find that parents have a low aversion for inequality in child outcomes. Because they perceive the returns to investments to be larger for children with higher initial conditions, they reinforce initial differences in their children's human capital. Household resources are also important, as constrained parents select more unequal allocations. I also show that primitive parameters identified in the experiment are predictive of the *actual* investments parents make in their children.

The results suggest that intra-household allocation of resources might be important to understand and measure inequalities across individuals in a society. At the same time they have important implications for the targeting of public policies as they suggest that a deeper understanding of intra-household dynamics might be important in order to design effective interventions that can work for *all* household members.

The final essay "Helping Struggling Students and Benefiting All: Peer Effects in Primary Education", joint with Samuel Berlinski and Matias Busso, studies peer effects in education. We use the randomized evaluation of a successful remedying education intervention that improved the academic achievement of low-achieving students within a class, to study spillover effects on their higher-achieving classmates. We find that the test scores of non-treated children in treatment schools increased by 0.108 of a standard deviation compared to similar children in control schools. We interpret this reduced-form effect on high-achieving students through the lens of a linear-in-means model of peer effects. We find that a one-standard-deviation increase in peers' contemporaneous achievement increases individual test scores by 0.679 of a standard deviation. We rule out alternative explanations coming from a reduction in class size, or a change in teachers' effort, and find suggestive evidence of a reduction in classroom disruption.

From a policy perspective, the findings suggest that programs aimed at improving the bottom of the achievement distribution have the potential to generate social-multiplier effects, so that it is possible to substantially improve academic outcomes for all with interventions targeted to the weakest. Moreover, our results provide the first successful example of how peer effects can be exploited in the design of public policies aimed at improving students' academic performance. We believe that these considerations are important to inform the policy debate concerned with the allocation of public funds to education.

Chapter 2

Child Development in the Early Years: Parental Investment and the Changing Dynamics of Different Dimensions

2.1 Introduction

It is well established that human capital constitutes an important factor of production and development. Fostering the process of human capital development can be crucial for growth and economic prosperity. From the point of view of poor families, the development of human capital can be an important factor in breaking the intergenerational transmission of poverty and poverty traps. The process through which individuals acquire different skills in the first part of the life cycle is very important in determining the level of inequality observed in a given society.

The characterization of the process of human development is therefore important. Over the last few decades, we have learned much about this process. For instance, we know that what happens in the early years is particular important for long term outcomes.[1] At the same time, it is now pretty much accepted that human capital is a multidimensional object. Its different domains are important both for the different roles they play in the process of development and as final outcomes, as they are related to different aspects of well-being and their combination determines the remuneration individuals get in the labour market.

However, there are still many aspects of human capital formation that are not fully understood. In particular, we still have an imprecise idea about the dynamic properties of such a process, and how they vary during the first few years of life. Moreover, the degree of persistence of various dimensions of human capital and the role they play in the growth process is still not characterized completely. Researchers

[1] See Currie & Almond (2011) for a review of the literature.

often assume that the process is of the Markov type, so that outcomes at age $a + 1$, after conditioning for outcomes at age a, do not depend on previous realizations of the process. While this assumption is often made for convenience and for the lack of accurate data, it is not an innocuous assumption. Deviations from a Markov process could explain, for instance, the fade-out in the impact of some intervention, followed by a subsequent re-emergence of the impacts. More generally, a flexible dynamic specification might be key in the identification of important ages in the process of development and *windows of opportunities* for specific interventions.

How different dimensions of development interact in the process of human capital formation and how they can be affected by external factors is also key for the design of appropriate policies and for the identification of the role played by different inputs in the process. A comprehensive characterization of these causal links might be of extreme importance to set the basis for the design of interventions that improve child development in deprived contexts. If one were to establish that certain skills can be influenced by specific factors at a particular age, and that those skills play an important role in the process of human capital formation, interventions that target those specific skills at that age should be the focus of policies that intend to foster child development.

In recent years, several studies have developed this research agenda. In a seminal paper, Cunha, Heckman, & Schennach (2010) specify a model of child development where different dimensions of human capital depend of past realizations of the process and some additional inputs, both observables and unobservables. Importantly, Cunha, Heckman, & Schennach (2010) allow the possibility that some variables of interest in the theoretical model are not observable. Researchers, however, have access to noisy signals of these latent variables of interest. Cunha, Heckman, & Schennach (2010) show the conditions under which such a model is non-parametrically identifiable. In their benchmark specification, they consider a flexible specification for the functional form that links the various components of the process and allow some inputs to be chosen by certain agents (such as parents) and therefore being related to unobservable variables.

The explicit recognition of the existence of measurement error proposed by Cunha, Heckman, & Schennach (2010) is important for several reasons. Such an approach spells out explicitly the assumptions about the available measures that permit to map observable variables into concepts that are relevant for the questions researchers are asking. It provides an effective way to use all the available measures to obtain efficient estimates of the parameters of interest. Furthermore, it allows to choose a metric for measurement, making the aggregate set of measures comparable over time and, possibly, across contexts.

Attanasio, Meghir, & Nix (2020) develop Cunha, Heckman, & Schennach (2010)'s contribution and explore alternative specification of the estimation procedure to analyse the development of children in India, using the Young Lives data. In addition to specifying a simpler and more flexible estimation procedure, they pay particular attention to the potential endogeneity of parental investment. Attanasio, Meghir, Nix, & Salvati (2017) also use the Young Lives data, but for Peru and Ethiopia. They use the innovations introcuced by Attanasio, Meghir, & Nix (2020) to explore the implication of alternative functional forms for the process of child development. Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020) apply the same methods and use the model they estimate, which includes the production function of human capital and parental behaviour, to interpret the evaluation findings of an early childhood intervention in Colombia.[2]

More recently, Agostinelli & Wiswall (2016a,b) have considered the issue of the metric to be used when modelling child development over different ages and stressed the risk of making strong assumptions in this respect. Repeated normalisations for the same factors at different ages may lead to important biases in the estimates of the exogenous growth process of child development (the equivalent of exogenous technical progress in a production function). In this paper, we address these normalization issue explicitly, as we are considering an extended time period of the early life cycle. As we discuss below, we have to be explicit about the normalization issues even because we do not have the same measures of child development available at all ages.

One of the reasons for the limited evidence on the dynamic properties of the process of human capital formation is the paucity of rich data that would allow researchers to identify the dynamic interactions among the various dimensions of human development, both during the early years and in subsequent periods. Moreover, one of the challenges researchers face even when longitudinal data are available, lies in the fact that available measures vary with children's age, so that the direct comparison and modelling interactions between different dimensions of development is not trivial. In this paper, we exploit the availability of a rich longitudinal data set that follows children every year from age 1 to age 7. At each age, several measures related to different dimensions of development are available. In what follows, we discuss the methodological problems we face to combine effectively this wealth of measures and how we tackle this problem.

This paper makes some original contributions to the literature that tries to characterize the process of child development. First, the fact that we have access to a set of high quality measures about several dimensions of child development covering a relatively long period of time at an annual frequency allows us to estimate processes of human capital formation in the early years that are much dynamically richer than those so far analysed in the literature. In particular, we are not forced to impose a

[2] Interestingly, they notice that, while the intervention was randomly allocated across a number of localities and evidently influenced parental investment, it cannot be used in their context to identify the causal link between parental investment and child development. Indeed that paper tries to identify the origin of the observed impact of the intervention and concludes that it has worked mainly by increasing parental investment. Posing this question cannot rule out that the intervention has a direct impact on child development.

Markov structure on the process, and we can allow for several lags to be relevant at different points in time.

Second, the fact that we have high quality information on several dimensions of child development allows us to study possible interactions among these dimensions and their effects in the process of human capital accumulation. Third, we also model the impacts of parental investment and estimate a rich and holistic process of child development that is of special utility in the design of early child development policies. A complete characterisation of the process allows us to identify *windows of opportunities* that can be of particular relevance for policy. The fact that we can model the impact of parental investment and establish its medium run effects tells us at what ages policies aimed at improving parental practices should be targeted. As we discuss below, this is relevant for at least two reasons: parental investment may be particularly effective at specific ages and, given its short run effectiveness, its medium run impacts depend on the dynamic features of the process. We show that it is better to improve parental investment when is most effective and just before the persistence of the process of development increases.

The paper also makes a small but important methodological contribution. Because for some particular skills we do not have a single measure available for all ages covered by our data, we need to link different measures over time. This involves some issues with the normalisation of the parameters of the measurement system we estimate below. Given the importance of the dynamics of the process for the results we discuss, these issues are particularly important.

The rest of this paper is structured as follows. Section 2.2 presents our model for the production of cognitive skills, socio-emotional skills, and health over the child's life-cycle, and describe how we deal with the endogeneity of parental investments. In that section, we also discuss in detail the dynamic properties of the process, paying particular attention to the issues of persistence and non-stationarity. Section 2.3 describes the latent factor approach that we use to aggregate the available information and take care of measurement error in the data. We pay particular attention to the issues related with the normalisation of the parameters of the measurement system, and propose a methodology to solve those issues. Section 2.4 describes the data and gives some background of the intervention. The main results are presented in Section 2.5, while the counterfactual exercises are presented in Section 2.6. Section 2.7 concludes.

2.2 The Production Function of Human Capital

In this section, we discuss our formulation of the dynamic process and sketch the way we model it. In particular, we start with a discussion of the dimensions we model. We then discuss the dynamics of the process and how we allow for a flexible specification. Finally, we discuss the role of different inputs and, in particular, that of parental investment. To do so, we use a simple model of investment which brings about the main assumptions that will drive our empirical strategy.

2.2.1 Different Dimensions: Cognition, Health and Socio-emotional Skills

As we mentioned above, it is well established that human capital is a multidimensional object. In the context we study, we decided to model the dynamic evolution of three different dimensions: cognitive skills, socio-emotional skills and health. Our choice is partly driven by the measurements available to us that are presumably related to such skills. Richer data sets might allow an even richer specification of the process. For instance, one could consider cognition separately from language development or, more importantly, one could model separately 'internalising" and "externalising" socio-emotional skills, as in Attanasio, Blundell, Conti, & Mason (2020). In any case, ours specification is one of the richest in the literature, especially with data from developing countries.[3]

As we discuss in Section 2.3, we map specific available measures to the three factors we model. The availability of detailed and high quality measures of development relevant for different measures is not easy, especially for socio-emotional skills. Having done that we can study the (dynamic) interactions among these different factors in the process of development. Such a structure is useful both from an academic and a policy point of view as it allows us to identify which specific skills might be more important at different ages in the process of child development. Such knowledge is possibly relevant for the design of specific interventions that might target one or another dimension of development.

2.2.2 The Dynamics of Human Development: Persistence and Non-stationarity

The other important and so far relatively unexplored features of the process of development are its dynamic properties. Often, again for data availability, researchers focus on models where the level of human capital at age a depends only on the level of human capital at age $a-1$ (and other variables). Such assumptions on the nature of persistence of the process are very strong and may be inconsistent with the evidence from certain interventions that show an effect in the short run that seems to disappear in the medium run and re-appear in the long run.[4] Moreover, very few data set give

[3] Attanasio, Meghir, & Nix (2020) and Attanasio, Meghir, Nix, & Salvati (2017), for instance, use the well-known Young Lives data to model simultaneously cognitive skills and health. Cunha, Heckman, & Schennach (2010) use data from the NSLY to model (at much older ages than those we consider here) cognition and socio-emotional skills. Rubio-Codina, Araujo, Attanasio, Munoz, & McGregor (2017) discuss some of the issues in measuring child development during the early years.

[4] See, for instance, Gertler et al. (2014) for the long run effect of an intervention in Jamaica, whose effects in the medium run had partly fade out, or Heckman, Moon, Pinto, Savelyev, & Yavitz

the possibility of studying how the persistence of the process changes through various ages. And yet these key parameters are crucial to determine the existence of *windows of opportunities* for child development and the optimal timing for policies aimed at improving the life chances of disadvantaged children.

In what follows, we consider flexible functional forms for the process of human development, allowing for two lags of the developmental process and letting the coefficients that determine persistence vary with age. The rich data set we have available, with several observations during the first few years, allows us to estimate these models. Armed with these estimates, we can evaluate the medium run effectiveness of different interventions targeting specific stages of the child's life cycle.

2.2.3 The Production Function of Human Capital

The model we have in mind for the process of human development can be represented by the following equation:

$$\boldsymbol{\theta_{t+1}} = f_t(\boldsymbol{\theta_t}, \boldsymbol{\theta_{t-1}}, \boldsymbol{X_t}, \boldsymbol{Z_t}, \boldsymbol{\epsilon_{t+1}}) \tag{2.1}$$

where the subscript t represents age. $\boldsymbol{\theta_t}$ is a vector which represent different dimensions of human development (e.g. cognition, health and socio-emotional skills), $\boldsymbol{X_t}$ and $\boldsymbol{Z_t}$ are vectors of (potentially) observable variables. The reason we differentiate them is that the variables in vector $\boldsymbol{X_t}$ are assumed to be chosen by some relevant agent, for instance parents, and could be reacting to the evolution of the process itself and would therefore be *endogenous*. The variables in $\boldsymbol{Z_t}$, instead, are environmental factors, including, parental background characteristics. The variables $\boldsymbol{Z_t}$ that we consider include family composition and parental background, such as education. The vector $\boldsymbol{\epsilon_{t+1}}$ represents unobservable (by the researcher) factors that affect the evolution of the various dimensions of human development.

The obvious endogenous variable $\boldsymbol{X_t}$ is parental investment. Parental behaviour and choices are recognised to be important determinants of child development.[5] However, they clearly must depend on the evolution of child development. The identification of their effect on child development is, therefore, difficult.

We keep the fuction f_t, which determines the evolution of human development over time, purposely vague, with the idea of wanting to use a flexible specification where different inputs may play an important role. We note the subscript t to the function f to stress the fact that we let it change with age. Indeed, one of main goals of this study is to establish how the function f changes over time.

The empirical characterisation of the function f_t in equation (2.1) faces different challenges. First, $\boldsymbol{\theta_t}, \boldsymbol{X_t}$ and $\boldsymbol{Z_t}$ are not observable directly or without measurement

(2010) for the long run impact of the well-known Perry Preschool Program.

[5] See, for instance, McCormick et al. (2020), Boonk, Gijselaers, Ritzen, & Brand-Gruwel (2018), Padilla & Ryan (2018), and Hsin (2007).

error. Second, while it is possible to estimate the main features of the process non-parametrically, in practice such an approach would require a very large data set, especially to characterize several dimensions of child development simultaneously, as we do here. Therefore, researchers often make functional form assumptions about f_t. It is therefore desirable to use parametric specifications that preserve a degree of flexibility while making the production function estimable. Finally and most importantly, as we mentioned above, some of the arguments of the function f_t are chosen in reaction to the evolution of human capital or simultaneously with other unobservable inputs. Such inputs are therefore *endogenous* and estimating their role in the production function is particularly hard. In this section, we discuss the second and third issue, while the discuss how we deal with measurement in Section 2.3.

Functional Form Assumptions

Given the data available, we model three factors, which we denote as cognitive skills (labelled C), socio-emotional skills (labelled S), and health (labelled H). For each of these factors and for each age considered, we estimate different functional forms for the production function. The first specification we consider is the Constant Elasticity of Substitution specification (CES). According to this specification, dimension k of child human capital at age $t+1$ is given by:

$$\begin{aligned}\theta_{k,t+1} = \; & A_{k,t}[\gamma_{k,C_1,t}\theta_{C,t}^{\rho_{k,t}} + \gamma_{k,C_2,t}\theta_{C,t-1}^{\rho_{k,t}} + \gamma_{k,S_1,t}\theta_{S,t}^{\rho_{k,t}} + \gamma_{k,S_2,t}\theta_{S,t-1}^{\rho_{k,t}} + \\ & \gamma_{k,H_1,t}\theta_{H,t}^{\rho_{k,t}} + \gamma_{k,H_2,t}\theta_{H,t-1}^{\rho_{k,t}} + \gamma_{k,I,t}\theta_{I,t}^{\rho_{k,t}} + \gamma_{k,P,t}\theta_{P}^{\rho_{k,t}}]^{1/\rho_{k,t}} \qquad k \in \{C,S,H\}\end{aligned} \tag{2.2}$$

where θ_t^I are parental investments, and θ^P represent parental skills, which are assumed to be fixed over time. A_t^k is a factor-neutral productivity parameter, which is allowed to depend on observable characteristics (such as family composition and gender), denoted by $\boldsymbol{G_t}$ and unobserved shocks according to:

$$A_{k,t} = exp(\delta_{k,0,t} + \boldsymbol{\delta'_{k,G,t}}\boldsymbol{G}_t + \epsilon_{k,t}) \tag{2.3}$$

Several comments on equation (2.2) are in order. Such a specification allows for a large number of inputs, including two lags of each of the dimension of human capital considered. The effect of these inputs on the total outcome is not additive nor separable. One can impose the assumption of constant return to scale by imposing that the sum of the γ coefficients is equal to 1, but it is not necessary. The CES allows for interactions, and the parameter $\rho_{k,t}$ governs the elasticity of substitution among the various inputs, which is given by $\frac{1}{1+\rho_{k,t}}$. When the parameter $\rho_{k,t}$ converges to zero, the expression in equation (2.2) converges to a Cobb-Douglas and the elasticity of substitution among the various inputs is unity. This writes as follows:

$$\theta_{k,t+1} = A_{k,t}[\theta_{C,t}^{\gamma_{k,C_1,t}}\theta_{C,t-1}^{\gamma_{k,C_2,t}}\theta_{S,t}^{\gamma_{k,S_1,t}}\theta_{S,t-1}^{\gamma_{k,S_2,t}}\theta_{H,t}^{\gamma_{k,H_1,t}}\theta_{H,t-1}^{\gamma_{k,H_2,t}}\theta_{I,t}^{\gamma_{k,I,t}}\theta_{P}^{\gamma_{k,P,t}}] \quad k \in \{C,S,H\} \tag{2.4}$$

All the arguments of the production function in these equations are potentially not observable directly. In Section 2.3, we discuss how to estimate the parameters of these functions when we have some noisy signals of these variables.

Although equation (2.2) is reasonably flexible (and contains as a special case the Cobb Douglas case), it does impose a fair amount of restrictions on the production function. In particular, the assumption that the degree of substitability among all the inputs is the same and governed by a single parameter ($\rho_{k,t}$) is a very strong one. For this reason we also experimented with a Transcendental Logarithmic production function (translog for short) which has both linear and quadratic terms and allows capturing different degrees of substitutability between inputs flexibly (Christensen, Jorgenson, & Lau (1973)). The translog specification (with some abuse of notation) is given by:

$$\begin{aligned} \ln\theta_{k,t+1} = \; & \gamma_{k,C_1,t}\ln\theta_{C,t} + \gamma_{k,C_2,t}\ln\theta_{C,t-1} + \gamma_{k,S_1,t}\ln\theta_{S,t} + \gamma_{k,S_2,t}\ln\theta_{S,t-1} \\ & +\gamma_{k,H_1,t}\ln\theta_t^H + \gamma_{k,H_2,t}\ln\theta_{H,t-1} + +\gamma_{k,I,t}\ln\theta_{I,t} + \gamma_{P,t}\ln\theta_P \\ & +\tfrac{1}{2}\textstyle\sum_q\sum_r \gamma_{k,qr,t}\ln\theta_{q,t}\ln\theta_{r,t} + \delta_{k,0,t} + \delta_{k,G,t}G_t + \epsilon_{k,t} \\ & k \in \{C,S,H\}; \quad q,r \in \{C,S,H,P,I\} \end{aligned} \tag{2.5}$$

with $\gamma_{i,j} = \gamma_{j,i}$. The translog production function reduces to the Cobb-Douglas when all $\gamma_{i,j} = 0$. Equation (2.5), as it is written, does not allow interactions with the lag-2 dimensions of human capital. In practice, we experiment with those interactions as well.

The main focus of the empirical analysis is going to be in assessing the extent of the persistence of the process, as measured by the coefficients on the lagged human capital variables, the extent of the dynamic interactions between different dimensions of human capital (for instance, what role socio-emotinal skills play in the development of cognition), and the role of parental investment. We will also assess how these effects change over time. As discussed above, the size of these parameters has important policy implications.

2.2.4 The Endogeneity of Parental Investment

Parents' choices reflect their resources (financial and time), their tastes and their beliefs. In making these choices they may react to the effect that some unobservable factors have on child development. Furthermore, not all parental choices are observable. Therefore the observed correlation between parental investment and children outcomes does not necessarily have a causal interpretation. In other words the estimates of the γ coefficients in equations (2.2), (2.4), or (2.5) obtained by simple OLS or NLS regressions can be biased. Indeed, the nature of the bias of such coefficients, if it can be estimated, is informative of the motives behind parental behaviour. For instance, a negative bias in the estimation of the marginal product of investment might be indicating compensatory behaviour, where parents react to negative shocks to their offspring's development by increasing investments.

To estimate the coefficients of the production function it is necessary to specify a model of parental behaviour, which can be used to derive the conditions under which the coefficients can be identified, and the variables that it is necessary to observe to obtain such estimates. In what follows, we assume that parents maximise a certain objective function subject to a resource constraint and their perception of the human capital production function. To clarify the main ideas, we sketch here a simple (and unrealistic) model of parental behaviour. In particular, we assume that couples choose investment to solve the following problem:

$$\begin{aligned} & Max_{c,X} U(c, H) \\ s.t. \quad & c + pX = Y \\ & H = g(X, H_0, W, \epsilon) \end{aligned} \tag{2.6}$$

where c is parental consumption, H child's human capital, X parental investment, p the price of investment in terms of consumption (whose price is normalised to 1), and Y represents parental financial resources. The function g is the parental perception of the human capital production function. The shock ϵ is observed by parents but not by the researchers estimating the parameters in (2.6).

In this simple model, we do not consider separately time and financial resources and we have a very simple utility function, where parental utility does not depend, for instance on X. Furthermore, we assume that parents observe the level of child development H_0, while in a more complex model, parents might have distorted beliefs about such a variable and make their choice on that basis. Such a simple model, however, is useful to discuss our empirical strategy. In such a model, parental investment depends on tastes, as represented by the utility function U, resources Y, the price of investment p and on the parental perceptions of the production function of human capital g. We note that the function g does not necessarily coincide with the "true" production function, as represented, for instance, in equation (2.2) or (2.5).

A feasible approach to estimate the parameters of the production function in this context would be the use of an IV or, in the non-linear case of equation (2.2), a control function approach. For such an approach to work, it is necessary to identify variables that drive investment, but that can be plausibly excluded from the production function. Once again, we note that, if the objective is only to estimate the parameters of the "true" production function, rather than the complete model, which would include taste parameters, the parameters of the "perceived" production function, one does not need to specify a completely accurate investment function or assume that parents know the "true" production function. Instead, one can estimate an investment function that includes some variables that drive investment and are not direct determinants of human capital.[6] This approach leads us to estimate the following approximation to the investment function:

$$\theta_{I,t} = \pi_0 + \pi_{C,t}\theta_{C,t} + \pi_{S,t}\theta_{S,t} + \pi_{H,t}\theta_{H,t} + \pi_{P,t}\theta_P + \pi'_{G,t}\boldsymbol{G_t} + \pi_{Z,t}\boldsymbol{W_t} + u_t \tag{2.7}$$

where the vector $\boldsymbol{G_t}$ are variables that affect the production function, while $\boldsymbol{W_t}$ are the instruments that do not enter the production function directly. Such an equation should be interpreted as a reasonable approximation (not necessarily consistent) of the investment function. Estimating it allows us to obtain estimates of u_t, which can then be used to construct a control function to add to the production function equation.

In the context of the model in equations (2.6), two natural candidate variables that could be used as instruments are the price of investment p and the financial resources Y. As discussed in Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020), prices could constitute a more attractive instrument than financial resources, as the former are plausibly taken as given by households, while the latter could proxy for some omitted input in the production function. Unfortunately, in the context that we will be analysing, there is no variation in prices, as the children in our sample live in two similar neighborhoods of the same city. We therefore decided to use variables related to financial resources as an instrument and discuss possible biases and interpretation of results in Section 2.5. Our results on the effectiveness of parental investment, therefore, should be taken with a grain of salt. It should be stressed that, in a similar context to the one considered here, Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020), who performed an extensive analysis of the instruments used to identify the effect of parental investment, obtain estimates of the relevant coefficient that do not depend on whether they use prices or resources as the source of identification and that are substantially different from the estimates obtained by OLS, which ignore the endogeneity of investment.[7]

2.3 Latent Factors and Measurement

As we mentioned above, several variables (or factors) that enter the production function (or the investment function) are not directly observed. Instead, our data consists of a set of measures that are related to the relevant latent factors. We follow the approach used, among others, by Cunha, Heckman, & Schennach (2010), Attanasio, Meghir, & Nix (2020) and Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020) to estimate a measurement system that allows us to synthetise the available

[6]In linear models, the consistency of estimates so obtained can be easily proven. The issue is subtler in non-linear model and one might have to show the robustness of the results with simulations.

[7]Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020) also discuss the possibility that the instruments they use are "weak" and therefore introducing a bias in the estimates. They show that instrument weakness is not an issue in their context.

information on these variables in a lower dimensional construct and account explicitly for the presence of measurement error. To obtain estimates of the relevant parameters governing the process of human capital formation, we need to address a number of normalisation and location issues. These issues can be particularly important in this context, as our focus is on the longitudinal and dynamic analysis of this process. As we discuss below, it is crucial to use, at each stage or age considered, a consistent metric and normalisation of the factors considered.

Using the approach introduced by Attanasio, Meghir, & Nix (2020), we obtain estimates of the joint distribution of the latent factors, which we then use in the estimation of the production functions we consider. We consider as factors three dimensions of human development (cognition and language, socio-emotional skills, and health) as well as parental investment and a set of indicators of parental background. We discuss the choice of these dimensions and the measurements we use in detail in Section 2.4.2.

In this section, we first discuss the measurement system and its estimation, including the normalisation issues we tackle. We then move on to consider the estimation of the production function, taking into account the process to obtain estimates of the relevant latent factors and the endogeneity of parental investment.

2.3.1 The Measurement System

We specify the measurement system in terms of the log of the latent factors, ensuring that the factors only take positive values, as required by the specification of the production function. Let $m^j_{k,t}$ be a continuous measure at age t for factor k.[8] Assuming a log-linear relationship between the measures and the factors with additive measurement error, we can write:

$$m^j_{k,t} = \alpha^j_{k,t} + \lambda^j_{k,t} \ln \theta_{k,t} + e^j_{k,t} \tag{2.8}$$

where $\alpha^j_{k,t}$ is an intercept term and $\lambda^j_{k,t}$ is a factor loading. The intercept captures the *difficulty* of measure $m^j_{k,t}$. The factor loading captures the *saliency* of variable j for factor k: a measure with a high value of $\lambda^j_{k,t}$ is one that varies much with variation in the factor. Both the intercept and the factor loading are indexed by t, reflecting the fact that these are allowed to vary by age to capture the potentially different informativeness of the same measure over time. The term $e^j_{k,t}$ represents the measurement error for variable j. This is assumed to be mean zero and independent of the latent factor. Furthermore, the error terms are assumed to be independent between each other ($e^j \perp\!\!\!\perp e^{j'}$) and over time ($e_t \perp\!\!\!\perp e_{t'}$).[9]

[8]If the available measure $m^j_{k,t}$ is discrete (binomial or otherwise), one can extend the discussion here considering a single index model such as a (multinomial) probit, a logistic or some other discrete variable model, which converts an equation such as (2.8) into a discrete measure.

[9]These assumptions can be somewhat relaxed, although we need to have some measures whose measurement errors are independent to achieve identification of the model's parameters.

The final and most important element in equation (2.8) is the latent factor $\theta_{k,t}$, which represents child development in dimension k at age t. Estimation of the measurement system in equation (2.8) allows us to learn the distribution of the child development factor in a population or sample. Moreover, for any set of measures observed for a given child, estimates of such a system allow us to construct an estimate of their development.

While, under certain conditions on the number of available measures, it is possible to identify the parameters in equation (2.8) without any assumptions on the distribution of the latent factor $\theta_{k,t}$, in practice researchers often make specific assumptions about such a distribution.[10] The parameters of the distribution of the latent factors constitute another set of parameters to be estimated. Finally, we notice that the way equation (2.8) is written, assumes that each measure correponds exclusively to one factor. Such a system is referred to as a *dedicated* measurement system. While it is possible to relax a bit this assumption, having some measures being affected by more than one factor, identification requires a set of exclusion restrictions.

Because the latent factors are not observable and have no natural metric, we need to impose some normalizing restrictions to set their location and scale (Anderson & Rubin (1956)). This issue is analogous to the discussion of *anchoring* in Cunha, Heckman, & Schennach (2010). As discussed by Agostinelli & Wiswall (2016a), normalizing the mean of the factor to be zero in each period can be a non-innocuous assumption when one is interested in capturing the dynamic growth of the factors over time, as it would restrict the set of admissible production functions to those that are mean stationary. Moreover, valid comparison of technology parameters over time, which is key for the identification of *windows of opportunities*, requires that the scale of the latent factors is consistently expressed in the same metric over ages. Expressing the scale of the latent factor arbitrarily, by normalizing the loading on a *different* measure in each period, would imply that the technology parameters are themselves expressed in different metrics over time.

One straightforward solution to both issues would be readily available if one had at least one same measure spanning all time periods. One could then express the location and scale of the factor in terms of this measure, which we label as measure 1, by imposing $\alpha^1_{k,t} = 0$ and $\lambda^1_{k,t} = 1 \; \forall t = 1...T$. The assumption that one of the intercepts $\alpha^h_{k,t}$ is normalised to 0 is necessary if we want to estimate, for each age, the mean of the latent factor $\theta_{k,t}$.[11] The assumption that the factor loading on one specific measure is set to 1 is a strong one and defines the scale of the estimated factors. This assumption says that the saliency of the measure whose loading factor is normalised to 1 does not change with age. As with the mean, this assumptions allows the variability

[10] Log-normality is often assumed. Alternatively researchers have used more flexible specifications, such as a mixture of log normal. See, for instance, the discussion in Cunha, Heckman, & Schennach (2010)

[11] Alternatively, some researchers leave the $\alpha^j_{k,t}$ unrestricted for every t and constraint the mean of the distribution of $\theta^j_{k,t}$ to zero. The two assumptions are observationally equivalent.

of the latent factors to change with age. Given these assumptions, we can estimate the measurement system in equation (2.8). Both the mean and the scale of estimated latent factor are now comparable over time as they are expressed in terms of this normalizing measure.

Unfortunately, while we do have one such measure in our data for health, socio-emotional skills and investments, we do not have it for cognitive skills. However, we can exploit the fact that there is at least one time period where we have some "overlapping" between different measures to express the location and scale of the cognitive factor consistently over time. In the next subsection we provide the details of this procedure with the help of an illustrative example.

Normalizations on the Location and Scale of the Latent Factors when no Measure is Available for all Ages

Our data consists of two sets of instruments that were used to measure child cognitive skills over time: the first set is only appropriate at younger ages (0 to 3), while the second set is only appropriate for older children (3+). Therefore the simple normalization discussed above cannot be implemented. We extend that procedure in a relatively simple way.

The basic intuition behind our procedure is the following: if there is at least one measure from each set that overlaps in at least in one time period then one can work out a mapping between the location and scale parameters of the two sets (the α_s and λ_s in equation (2.8)) and impose these restrictions in subsequent periods, so to set the location and scale of the latent factors in a common metric over time.

The approach is explained simply with an example where we consider three time periods and four measures for a generic factor k: $m_k^{a_1}$, $m_k^{a_2}$,$m_k^{b_1}$, $m_k^{b_2}$. To explain our procedure, we only need one latent factor. For simplicity, therefore, we omit the measurement equations relating to factors different from k from the discussion and drop the subscript k in what follows.

We assume that at age $t = 1$ m^{a_1} and m^{a_2} are available. At age $t = 2$, measures m^{a_1} and m^{b_1} are available, while at age $t = 3$, m^{b_1} and m^{b_2} are available. At this point the distribution of the (log) latent factor is left unspecified with mean μ_θ and variance σ_θ^2. With these assumption, the measurement system can be expressed as follows:

$$t = 1 \quad \begin{cases} m_1^{a_1} = \alpha_1^{a_1} + \lambda_1^{a_1} \ln \theta_1 + e_1^{a_1} \\ m_1^{a_2} = \alpha_1^{a_2} + \lambda_1^{a_2} \ln \theta_1 + e_1^{a_2} \end{cases}$$

$$t = 2 \quad \begin{cases} m_2^{a_1} = \alpha_2^{a_1} + \lambda_2^{a_1} \ln \theta_2 + e_2^{a_1} \\ m_2^{b_1} = \alpha_2^{b_1} + \lambda_2^{b_1} \ln \theta_2 + e_2^{b_1} \end{cases} \tag{2.9}$$

$$t = 3 \quad \begin{cases} m_3^{b_1} = \alpha_3^{b_1} + \lambda_3^{b_1} \ln \theta_3 + e_3^{b_1} \\ m_3^{b_2} = \alpha_3^{b_2} + \lambda_3^{b_2} \ln \theta_3 + e_3^{b_2} \end{cases}$$

In this example, no measure is available throughout the 3 periods considered. However, both measure m^{a_1} and measure m^{b_1} give a way to construct a metric for the factors that can be used through the 3 periods. In particular, if we normalise the intercepts and factor loadings of measures m^{a_1} to 0 and 1 respectively in each time period ($\alpha_1^{a_1} = \alpha_2^{a_1} = 0$ and $\lambda_1^{a_1} = \lambda_2^{a_1} = 1$), we can use the estimates of the intercept and factor loading of $m_2^{b_1}$ to express the location and scale of the latent factor at age 3 in same metric. We can do so by imposing that the location and scale coefficients pertaining to this measure are constant across time periods. Alternatively, a similar exercise could be performed normalizing the parameters of measure m^{b_1} and use that measure as the metric for the three periods. We notice that the parameters of the joint distribution of factors (μ_θ and σ_θ^2) can be identified. Our procedure guarantees that these are constantly expressed in the same metric and are, therefore, comparable over time. More formally, in our example, we would use the following procedure:

Factor scale We can set the scale of the (log) factor in the first two periods to be equal to each other by setting $\lambda_1^{a_1} = \lambda_2^{a_1} = 1$. In this way the scale of the latent factor would be the same in the first two periods and would be given by the scale of measure m^{a_1}. As the factor loading on measure m^{a_1} in period $t = 2$ is normalized to one, the factor loading on measure m^{b_1} in that period will be expressed in terms of measure m^{a_1}.[12] We can then impose $\lambda_2^{b_1} = \lambda_3^{b_1}$ to express the scale of the factor in $t = 3$ to be the same as that of the factor in $t = 1, 2$.

Factor location We can express the location of the (log) factor in $t = 1$ and $t = 2$ by setting $\alpha_1^{a_1} = \alpha_2^{a_1} = 0$. Taking expectations of the first equation at $t = 1$ and $t = 2$ we get $E(\ln \theta_1) = E(m_1^{a_1})$ and $E(\ln \theta_2) = E(m_2^{a_1})$, respectively. In this way the factor mean in periods 1 and 2 is expressed in terms of the same measure. Using the measurement equation for $m_2^{b_1}$, we can get an expression for $\alpha_2^{b_1}$:

$$\alpha_2^{b_1} = E(m_2^{b_1}) - \lambda_2^{b_1} E(\ln \theta_2) \tag{2.10}$$

[12] This approach is analogous to express the value of a currency in terms of another i.e. a nominal exchange rate.

Using the measurement equation for $m_3^{b_1}$, we can then get an expression for $E(\ln\theta_3)$:

$$E(\ln\theta_3) = \frac{E(m_3^{b_1}) - \alpha_3^{b_1}}{\lambda_3^{b_1}} \tag{2.11}$$

Finally, imposing $\alpha_2^{b_1} = \alpha_3^{b_1}$ and $\lambda_2^{b_1} = \lambda_3^{b_1}$ and substituting (2.10), we get:

$$E(\ln\theta_3) = E(\ln\theta_2) + \frac{E(m_3^{b_1}) - E(m_2^{b_1})}{\lambda_2^{b_1}} \tag{2.12}$$

so that the mean of the (log) factor in $t = 3$ will be equal to that of the (log) factor in the previous period (which was expressed in terms of measure m^{a_2}), plus a term that depends on the change in measure m^{b_1} between periods 2 and 3, scaled by the informativeness of this measure in terms of m^{a_2}.

This example can be easily generalised to additional time periods and additional available measures. What we want to make clear is that the availability of two measures in a given time period allows us to "convert" one to the other. We can then impose these conversions in different periods and express the scale and location of the latent factors in a way that is consistent throughout the periods considered. The resulting factors can then be compared over time to establish the rate of growth over time and to derive estimates of the parameters governing the growth process. The crucial identification assumption is that the factor loadings and the intercepts of some of the measures available at different ages are constant over time. In what follows, we will be using the *Test de Vocabulario en Imàgenes Peabody* (Dunn, Padilla, Lugo, & Dunn (1986)) to express the scale and the location of the cognitive factor.

2.3.2 Estimation

In order not to impose strong restrictions on the technology of human capital formation, one needs to estimate the distribution of factors using a flexible specification. Often used assumptions, such as joint normality, in practice impose very strong restrictions on the production function. For instance, the linear relationship among conditional means implied by joint normality would preclude any complementarity between the inputs in the production function. Thus flexibility in the specification of the production technology requires flexibility in the parametric assumptions about the joint distribution of factors.

In this paper, we follow the approach introduced by Attanasio, Meghir, & Nix (2020). This approach consists in assuming some flexible distributional assumptions for the latent factors and measurement errors, and estimating the parameters of these distributions (together with the parameters of the measurement system) using the observable data. In practice, we assume that the factors are distributed as a mixture of two log-normal distributions and that the measurement errors are normally distributed with mean zero and diagonal variance-covariance matrix. These distributional assumptions, together with the specified measurement model, imply that the measures will be distributed as a mixture of normal distributions. We thus estimate the parameters of the joint distribution of measurements by maximum likelihood, using the EM algorithm. Using these estimates we can recover the parameters of the joint distribution of factors, the variances of measurement errors, the factor loadings and the intercepts by minimum distance. Following Attanasio, Meghir, & Nix (2020), we augment the measurement system for the latent factors reflecting child development and parental investment, with additional equations, representing the control variables we use as well as the instruments used in the investment equation to take into account the endogeneity of such a variable.

Having obtained estimates of the joint distribution of all the variables included in the model, we use it to draw vectors of random variables from this distribution. As we estimate the joint distribution, we reproduce the correlation oberved in the data and use these synthetic data to obtain parameters of the structural model using a control function approach that we discuss in what follows.

2.3.3 Production Functions and Investment Functions: a Control Function Approach

An additional challenge we face in the estimation of a production function, is the fact that some of the arguments are endogenous. In particular, parental investment is likely to be chosen by parents in reaction to the evolution of child development. Moreover, the measured factor might be correlated with other unobservable factors that are relevant for child development. These issues make the identification of the causal link between parental investment and child development challenging.

We use the approach in Attanasio, Meghir, & Nix (2020), by first estimating an *investment function*, which represents the reduced form of a model that relates parental investment to (some of) its determinants. We then use the residual terms of such an equation to construct a *control function*, which can be used to control for the endogeneity of investment in the production function of human capital. The identification assumption is that the investment function includes some variables that are important drivers of investment but that do not enter directly the production function. Both the inverstment equation, its residuals and the production function augmented with the relevant control function are estimated on the synthetic data drawn from the joint distribution derived from the measurement system.

2.4 *aeioTU*: Available Data and Latent Factors

The longitudinal data set we use in this paper comes from a randomized controlled trial used to evaluate a comprehensive, high-quality early education intervention in Colombia. We use a sample of 1,073 disadvantaged children aged zero to four at baseline in two communities in northern Colombia. Bernal, Giannola, & Nores (2020) and Nores, Bernal, & Barnett (2019) provide additional details of the intervention and its impacts on child development.

2.4.1 Background on the Intervention and Its Evaluation

The data we use were collected within the evaluation of an important intervention aimed at improving the quality of preschool attended by disadvantaged children in Colombia. In particular, the intervention established two large child care centres, while at the same time implementing a specific curriculum used within them. The curriculum, inspired by the celebrated *Reggio Emilia* model, is promoted by an NGO, called *aeioTU*, which through a public-partnership with the Colombian government, operated 28 centres by 2016, providing comprehensive early childhood education to about 13,300 low-income children aged zero to five throughout the country. The program offers early education on a full-day schedule, five days a week and 11 months per year, plus nutritional supplementation corresponding to 70% of childrens daily calorie intake requirements.

The evaluation of the intervention was designed as a randomized control trial, which exploited the excess demand of places in two renewed child care centres on the Caribbean coast in Colombia. The RCT study assigned the children of families who applied to the two centres included in the study to treatment or control. The two centres in the study had capacity for 320 children, with just over half of that for children up to age three. The randomisation stratified children by age (five groups), gender and neighborhood within site (three groups). Slots were randomized for 1,073 children aged zero to four. Of these children, 471 were offered slots in the centers, and 602 children were allocated to the control group.

Baseline assessments were conducted on children in late 2010, prior to random assignment and to the beginning of the intervention. Children and their families were assessed every subsequent year (roughly every 10 to 12 months) up until 2014. Baseline data collection in Site 1 took place on July-September 2010 (Y1 henceforth) and the program started in November 2010. Baseline in Site 2 was conducted on October-December 2010 with the program starting in March 2011. The first follow-up (Y2) was collected 8 months into the program, a second follow up (Y3) 20-22 months after the start of the program, a third follow-up (Y4) at 32 months, and the last follow-up (Y5) at approximately 41 months since the start of the intervention.

2.4.2 Available Measures: Different Measures and Different Ages

The data we use is very rich, in that for each of the latent factors that enter the production functions we estimate, we have several measures available. We use the

Table 2.1: Measures of Child Development by Developmental Domain and Study Wave

	Y1	Y2	Y3	Y4	Y5
Health and Nutrition					
Height, weight, arm circumference	0+	0+	0+	0+	0+
Cognitive Development					
Bayley 3rd edition	0-42	0-42	0-42	0-42	-
Peabody Picture Vocabulary Test (Spanish)	30+	30+	30+	30+	30+
ELSA Early Literacy Skills Assessment	36+	36+	36+	36+	-
Woodcock-Muñoz III subscale 10[a]	36+	36+	36+	36+	36+
Woodcock-Muñoz III (subscales 1[b] and 9[c])	-	-	-	36+	36+
Head Toes Knees and Shoulders (HTKS)	48+	48+	48+	48+	48+
Socio-emotional Development					
Ages & Stages socio-emotional domain	6-60	6-60	6-60	6-60	6-60
Behavioral and emotional screening system	-	-	36+	36+	36+
Vineland Adaptive Behavior Scales-II	-	-	-	36+	36+

Notes: Each cell reports the ages for which the measure is available (in months).[a] Applied problems, [b] Word identification, [c] Text comprehension.

factor models described in Section 2.3 to summarize all the information they provide. In particular, the available data, an exploratory factor analysis and *a-priori* considerations lead us to consider three latent factors of child development: cognition and language, socio-emotional skills, and health.

The instruments used in the evaluation have adequate psychometrics and been used extensively in evaluations of early care and education including studies in developing countries (Fernald, Prado, Kariger, & Raikes (2017)). Child assessments were collected by graduates in psychology and students in their senior year in psychology, who were trained to reliability standards (100% agreement with the trainer) by experienced staff in a two-week training which included live reliability with young children.[13] Data collection was conducted in spaces rented and adapted for that purpose every year, under identical conditions for treatment and control children, with parental informed consent. Parent interviews were carried out in a separate room alongside the child's assessment.

Table 2.1 summarizes the child assessments used in the evaluation, by developmental domain. These changed as children aged over the course of the study. For health outcomes, we collected height, weight, and arm circumference following World Health Organization (WHO) standards (WHO (2007)).

Cognitive development was measured using the cognitive, motor, and language scales from the Bayley Scales of Infant Development III (BSID) (Bayley (2005)), for children under 36 months of age. We administered the Test de Vocabulario en Imágenes Peabody (TVIP) (Padilla, Lugo, & Dunn (1986)) which measures receptive language to children over 30 months of age in each wave. We used the Early Literacy Skills Assessment (ELSA) (DeBruin-Parecki (2005)) to measure early literacy in children over 36 months of age in each wave except the last. Math and literacy skills

[13] Evaluators involved in child assessment at any given wave were offered a refresher training every year, and new assessors (if any) were fully trained in similar conditions.

were measured using three subtests of the Woodcock-Muñoz III Tests of Achievement (WM-III): subtests #1 (letter-word identification), #9 (text comprehension) and #10 (applied problems) (Muñoz-Sandoval, Woodcock, McGrew, & Mather (2005)). The applied problems subtest was used every year for children above three years of age, while the literacy sub-tests were included only starting in Y4. Executive Function was assessed using the Head-Toes-Knees and Shoulders (HTKS) which examines behavioural regulation in young children (Ponitz Cameron et al. (2008), Ponitz, McClelland, Matthews, & Morrison (2009)). HTKS requires children to remember, respond and inihibit behavioural commands. We measured HTKS for all children older than four in all waves.

Socio-emotional development was assessed using The Ages and Stages Questionnaire: Socio-Emotional (ASQ:SE), a parent-completed assessment for children 6-60 months old measuring self-regulation, compliance, communication, adaptive functioning, autonomy, affection, and interactions with others (Squires, Bricker, & Twombly (2002)). It was collected in all waves. We also used the Vineland Adaptive Behavior Scales (Sparrow, Cicchetti, & Balla (1984)) on children above age three starting Y4. The Vineland is a parent questionnaire on personal and social skills, daily living skills, socialization, and motor skills. As children outgrew the ASQ:SE (starting at Y3), we used the Behavior Assessment System for Children, Second Edition (BASC-II) for all children older than 36 months of age, which measures adaptive functioning and problem behaviours (Bracken, Keith, & Walker (1998), Doyle, Ostrander, Skare, Crosby, & August (1997)).

Finally, we measure parental investments using items adapted from the Family Care Indicators (FCI) developed by UNICEF (Frongillo, Sywulka, & Kariger (2003)). This instrument was used to collect information about the types and numbers of play materials available around the house and about the types and frequency of play activities the child engages in with an adult.

In addition to the outcome measures described above, we included a household survey (primarily answered by mothers or the head of the household) inquiring on schooling attainment, maternal age at birth of the child, race, income and expenditures, employment, assets, health insurance, number of children in the household, and childcare experiences.

2.4.3 Impacts on Child Outcomes

Bernal, Giannola, & Nores (2020) reports positive effects on health throughout the study which were concentrated on boys, and positive effects on cognition for girls but not for boys. No significant program impacts were found on socio-emotional development. The effects on health were around 0.12 standard deviations throughout the study. The effects on cognition started at 0.16 standard deviations at the first follow-up (8 months into the program) and reached 0.36 standard deviations after that. The effects are observed from Y2 to Y4, but disappear in Y5. The treatment on the treated effect was larger with an effect of about 0.3 standard deviations for each year of effective enrolment in the program. These effects can be thought as percentages of the socio-economic development gap in Colombia, which is close to one standard deviation in receptive vocabulary at age five (Bernal, Martínez, & Quintero (2015)). For a detailed discussion on the evaluation results see Bernal, Giannola, & Nores (2020) and Nores, Bernal, & Barnett (2019).

Of course, a finer decomposition of child development and a larger number of factors could be possible. For instance, one could consider different dimensions of cognitive development, including language, executive functions and other types of cognitive skills. Or one could decompose what we call socio-emotional skills into "internalizing skills" (focus, motivation, drive and grit) and "externalising skills" (sociability, the ability to work and communicate with others). However, given the data available and the exploratory analysis we performed, the decomposition we perform seems reasonable and feasible. These three factors, have been used in the literature on child development extensively, although we are the first to model the three of them jointly.

2.4.4 The Measurement System

Given the available measures, we specify the measurement system as a *dedicated* system. In Table 2.2, we report the observable variables that we use as markers for each of the unobservable factors we include (cognition, socio-emotional skills, health, and parental investment). After estimating the measurement system for each factor and each age, following the normalization and scaling procedures discussed above, we can estimate, for each of the measures that we include in the system, the fraction of the variance of that variable which is accounted for by variation in the relevant unobservable factor and the fraction that is due to measurement error. These results are reported in Table 2.2.

Starting with the cognitive measures, we notice that they change with the age of the children, as is normal practice. Fortunately, there is some overlap at each age, so that we can follow the normalization and anchoring strategy we discuss above. At early ages, from one to three, the sample children are administered various sub-scales of the Bayley Scale of Infant Development, a test that is considered the gold standard for these ages. The quality of the Bayley tests is reflected in the high signal to noise ratios corresponding to these measures that we report in Table 2.2. We notice in particular, the high signal to noise ratio for the cognitive and the two language scales (expressive and receptive). Starting age three, children are also administered the TVIP, which we use as our metric for the cognitive factor. From age four, children are also administered the Woodcock- Munoz test and the ELSA test. While these measures perform reasonably well, the signal to noise ratio is not as high as that of the Bayley or the TVIP.

Table 2.2: Share of Total Variance Due to Signal and Noise

	% signal	% noise		% signal	% noise
Child cognitive skills			*Child socio-emotional skills*		
Bayley cognition age 1	0.91	0.09	ASQ selfregulation age 6	0.4	0.6
Bayley expressive age 1	0.58	0.42	ASQ communication age 6	0.28	0.72
Bayley receptive age 1	0.61	0.39	ASQ adaptive functioning age 6	0.17	0.83
Bayley fine motor age 1	0.86	0.14	ASQ affection age 6	0.33	0.67
Bayley gross motor age 1	0.9	0.1	ASQ compliance age 6	0.4	0.6
Bayley cognition age 2	1	0	ASQ interaction age 6	0.29	0.71
Bayley expressive age 2	0.54	0.46	*Child health*		
Bayley receptive age 2	0.44	0.56	Weight age 1	0.67	0.33
Bayley fine motor age 2	0.6	0.4	Arm circumference age 1	0.24	0.76
Bayley gross motor age 2	0.6	0.4	Height age 1	0.79	0.21
TVIP age 3	1	0	Weight age 2	0.71	0.29
Bayley cognition age 3	0.89	0.11	Arm circumference age 2	0.09	0.91
Bayley expressive age 3	0.6	0.4	Height age 2	0.69	0.31
Bayley receptive age 3	0.38	0.62	Weight age 3	0.88	0.12
Bayley fine motor age 3	0.5	0.5	Arm circumference age 3	0.1	0.9
Bayley gross motor age 3	0.43	0.57	Height age 3	0.91	0.09
TVIP age 4	0.71	0.29	Weight age 4	0.73	0.27
Woodcock-Munoz age 4	0.26	0.74	Arm circumference age 4	0.25	0.75
ELSA age 4	0.38	0.62	Height age 4	0.9	0.1
TVIP age 5	0.27	0.73	Weight age 5	0.67	0.33
Woodcock-Munoz age 5	0.32	0.68	Arm circumference age 5	0.24	0.76
ELSA age 5	0.36	0.64	Height age 5	0.98	0.02
TVIP age 6	0.43	0.57	Weight age 6	0.81	0.19
Woodcock-Munoz age 6	0.54	0.46	Arm circumference age 6	0.51	0.49
ELSA age 6	0.23	0.77	Height age 6	0.67	0.33
TVIP age 7	0.42	0.58	Weight age 7	0.79	0.21
Woodcock-Munoz age 7	0.44	0.56	Arm circumference age 7	0.55	0.45
ELSA age 7	0.15	0.85	Height age 7	0.64	0.36
Child socio-emotional skills			*Parental investments*		
ASQ selfregulation age 1	0.12	0.88	Child books ages 1-2	0.05	0.95
ASQ communication age 1	0.01	0.99	Number of play materials ages 1-2	0.41	0.59
ASQ adaptive functioning age 1	0.22	0.78	Books at home ages 1-2	0.93	0.07
ASQ compliance age 1	0.63	0.37	Toys ages 1-2	0.08	0.92
ASQ interaction age 1	0.01	0.99	At least 3 child books ages 1-2	0.08	0.92
ASQ selfregulation age 2	0.11	0.89	Learning materials (score) ages 1-2	0.07	0.93
ASQ communication age 2	0.15	0.85	Variety (score) ages 1-2	0	1
ASQ adaptive functioning age 2	0.02	0.98	Child books ages 3-4	0.15	0.85
ASQ compliance age 2	0.24	0.76	Books at home ages 3-4	0.06	0.94
ASQ interactiona age 2	0.09	0.91	Number of play materials ages 3-4	0.84	0.16
ASQ selfregulation age 3	0.18	0.82	Toys to learn colors ages 3-4	0.73	0.27
ASQ communication age 3	0.2	0.8	Games ages 3-4	0.52	0.48
ASQ adaptive functioning age 3	0.19	0.81	Soft toys ages 3-4	0.69	0.31
ASQ affection age 3	0.16	0.84	Toys to learn numbers ages 3-4	0.59	0.41
ASQ compliance age 3	0.34	0.66	At least 10 child books ages 3-4	0.73	0.27
ASQ interaction age 3	0.11	0.89	Mother activities with child ages 3-4	0.08	0.92
ASQ selfregulation age 4	0.2	0.8	Child books ages 5-6	0.13	0.87
ASQ communication age 4	0.26	0.74	Books at home ages 5-6	0.05	0.95
ASQ adaptive functioning age 4	0.15	0.85	Number of play materials ages 5-6	0.82	0.18
ASQ affection age 4	0.5	0.5	Toys to learn shapes ages 5-6	0.57	0.43
ASQ compliance age 4	0.25	0.75	Puzzles ages 5-6	0.48	0.52
ASQ interaction age 4	0.3	0.7	Toys for music ages 5-6	0.07	0.93
ASQ selfregulation age 5	0.15	0.85	Games ages 5-6	0.6	0.4
ASQ communication age 5	0.64	0.36	Soft toys ages 5-6	0.62	0.38
ASQ adaptive functioning age 5	0.14	0.86	Toys to learn numbers ages 5-6	0.66	0.34
ASQ affection age 5	0.42	0.58	At least 10 child books ages 5-6	0.28	0.72
ASQ compliance age 5	0.23	0.77	Toys to learn names ages 5-6	0.46	0.54
ASQ interaction age 5	0.34	0.66	Musical instruments ages 5-6	0.48	0.52

Notes: †0 and 1 reflect rounding

Table 2.3: Evolution of the Latent Factors Standard Deviations with Age

	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
Cognitive factor	0.55	0.36	0.26	0.29	0.41	0.60	0.60
Socio-emotional factor	0.45	0.35	0.43	0.51	0.47	0.65	-
Health factor	0.39	0.29	0.32	0.34	0.39	0.58	0.73

For socio-emotional skills, children are administered several ASQ sub-scales. Unfortunately, the signal to noise ratio for several of these sub scales is relatively low. The relative low information provided by these measures is one of the reasons that prevents us from having different socio-emotional factors.

For health, the three measures we have are weight, height and arm circumference. It is interesting to note that in the first few years of life, weight and height are very informative, while arm circumference starts to be somewhat informative at age four and becomes more important past age five.

Finally, for parental investment we have several measures from the widely used *Family Care Indicators* (FCI) test. As we can see from the table, several of these measures are good markers of parental investment and contain a substantial amount of information. However, we stress that most items we use refer to material (rather than time) investment, for which we had information of very limited quality.

2.4.5 The Evolution of Child Development Over Time: Descriptive Evidence

As we estimate the measurement system for the three latent factors we consider, and for all the ages in our sample, we effectively estimate the distribution of the latent factors, and how these distributions change with age. In Figure 2.1, we plot the means of the cognitive, socio-emotional and health factors, as implied by the estimated measurement system.

We observe that the mean of the cognitive factor, as to be expected, grows with age. The mean of the socio-emotional factor, however, is relatively flat: after a dip after age one, it recovers after age four, but the observed changes are not major. Finally, for health, the estimated means grow monotonically with age.

The means of the three factors of child development are not the only interesting moments to consider. Having estimated the entire distribution for our sample and how that distribution changes with age, it is interesting to consider this piece of evidence. In Figure 2.2, we plot the evolution of the distribution for the three factors we estimate over age, while, in Table 2.3, we report how the standard deviations of these factors change with age. We notice that, for the three factors and for all the ages considered, the estimated distributions are uni-modal. It is worth noticing that this feature is not imposed in estimation, as we assume that the distribution of the factors is given by a mixture of normal distributions. We also notice that, for the three factors, the dispersion of the distribution increases considerably at age 6. Interestingly, for the cognitive factor, the distribution at age 3 and 4 is considerably tighter.

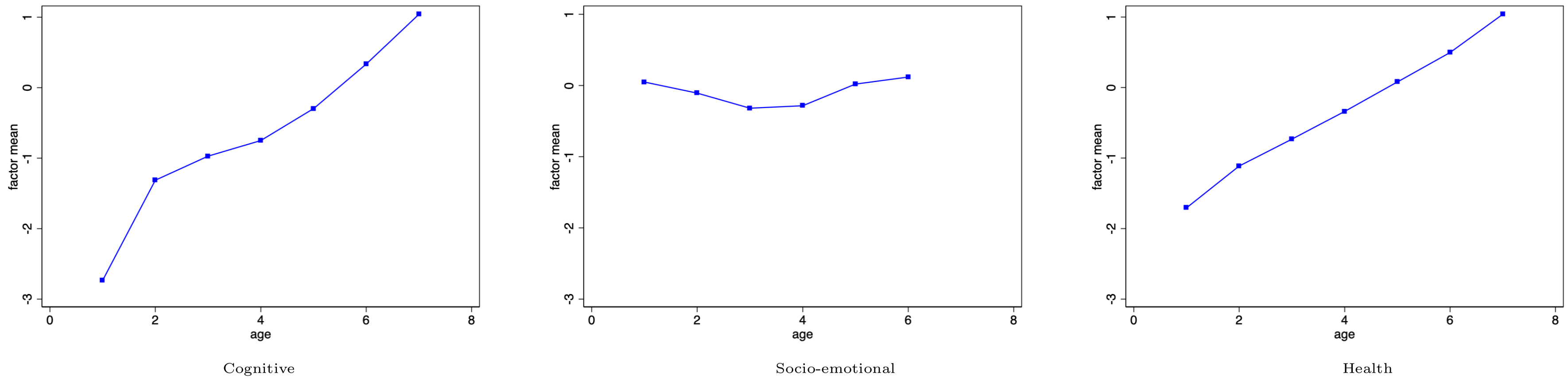


Figure 2.1: Latent Factors' Means Over Time

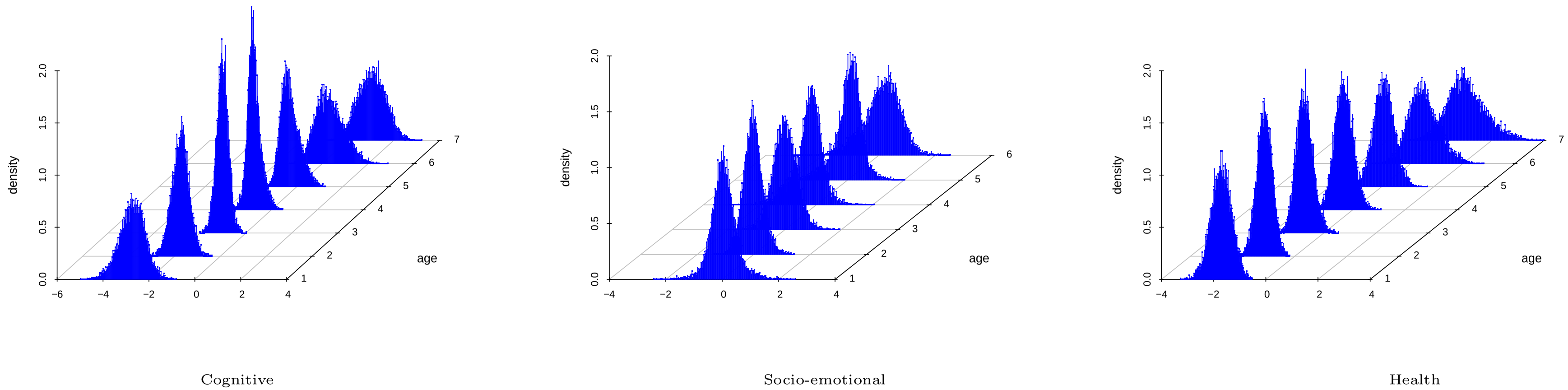


Figure 2.2: Latent Factors' Distributions Over Time

2.5 Parental Investment and the Dynamics of Child Development

Having established what are the latent factors that we use in our empirical exercise, and having estimated a measurement system, we can proceed to use estimates of the distribution of the relevant factors from the available measures to estimate an investment function and a *production function* for the various dimensions of child development. This production function establishes the role played by parental investment and other factors in the process of development. Furthermore, it characterises the dynamics of the process and the interactions of different factors at different ages.

2.5.1 Investment Functions

The first use we make of the factor estimates obtained from the measurement system is to estimate the investment functions (2.7) discussed in section 2.2.4. In particular, we assume that at each age, parental investment is a function of the developmental status of the children (as measured by the three factors we consider), of parental cognition, the number of children in the house, the gender of the child, as well as income, household size, and an indicator for the presence of the father. When interpreting these coefficients we should stress that this is (mostly) material investment in the home, and therefore it excludes time investment as well as investments that parents might perceive their children receive in the nursery or school they attend. The results are reported in Table 2.4.

Starting with the reaction of parents to their children's conditions, we find that the coefficient on cognition in the investment equations is not significantly different from zero at ages 2 , 3 and 5 and it is marginally significant and negative at age 4. This would seem to indicate that children with a lower cognitive development receive a bit more investment at age 4. The coefficients at age 6 and 7, however, identify a strongly significant and positive effect. It seems that at those ages, better developed children receive higher investment.

The coefficients on current socio-emotional skills are never statistically different from zero, with the marginal exception of age 6, where the effect is positive but only significant at the 10% level. Finally, the coefficient on health is marginally statistically positive at age 5 and is not different from zero at all other ages.

Parental cognition has a strong positive effect on investment, which is statistically significant at all ages. The coefficient increases with age and is particularly large at ages 6 and 7. The number of children in the household has a negative impact, which is significantly less then zero at all ages until 5. We do not find any statistically significant effects of child gender on parental investment.

The presence of the father has a strong positive effect on parental investment, while total household size has a negative one. The coefficient on income is always

Table 2.4: Investment Equations

	Investment equations					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
Intercept	0.11	0.071	-0.264	-0.103	-0.155	-0.266
	0.206	0.165	0.097	0.103	0.085	0.077
	[-0.254,0.417]	[-0.221,0.327]	[-0.394,-0.071]	[-0.201,0.115]	[-0.252,0.028]	[-0.352,-0.092]
Cognition t	0.084	0.026	-0.099	0.115	0.147	0.093
	0.078	0.087	0.054	0.091	0.077	0.062
	[-0.045,0.205]	[-0.081,0.202]	[-0.198,-0.01]	[-0.014,0.268]	[0.035,0.279]	[-0.01,0.179]
Socio-emotional t	-0.231	-0.05	-0.027	-0.01	0.11	0.022
	0.29	0.114	0.036	0.051	0.07	0.049
	[-0.392,0.491]	[-0.163,0.199]	[-0.088,0.023]	[-0.071,0.088]	[0.027,0.25]	[-0.039,0.123]
Health t	0.045	0.201	0.085	0.118	0.018	0.043
	0.125	0.103	0.067	0.066	0.054	0.04
	[-0.155,0.248]	[-0.024,0.302]	[0,0.217]	[-0.001,0.216]	[-0.072,0.114]	[-0.016,0.118]
Parental Cognition	0.062	0.059	0.233	0.229	0.35	0.357
	0.125	0.114	0.067	0.064	0.092	0.095
	[0.011,0.405]	[0.007,0.377]	[0.144,0.345]	[0.14,0.341]	[0.191,0.489]	[0.192,0.493]
Number of Children	-0.259	-0.296	-0.314	-0.317	-0.329	-0.347
	0.115	0.083	0.058	0.059	0.072	0.071
	[-0.445,-0.057]	[-0.404,-0.143]	[-0.391,-0.197]	[-0.375,-0.177]	[-0.395,-0.157]	[-0.409,-0.176]
Gender	0	0.003	0.043	0.035	0.033	0.029
	0.034	0.031	0.026	0.026	0.031	0.032
	[-0.055,0.053]	[-0.043,0.055]	[-0.01,0.077]	[-0.015,0.069]	[-0.029,0.081]	[-0.04,0.072]
Income	0.185	0.197	0.233	0.232	0.175	0.179
	0.067	0.061	0.044	0.043	0.043	0.044
	[0.064,0.274]	[0.075,0.268]	[0.15,0.29]	[0.14,0.277]	[0.082,0.227]	[0.085,0.232]
Household size	-0.293	-0.295	-0.106	-0.091	-0.056	-0.063
	0.077	0.065	0.044	0.043	0.067	0.067
	[-0.414,-0.168]	[-0.39,-0.184]	[-0.166,-0.015]	[-0.146,-0.002]	[-0.156,0.057]	[-0.161,0.052]
Father present	0.149	0.145	0.126	0.135	0.104	0.12
	0.04	0.04	0.023	0.024	0.03	0.028
	[0.082,0.211]	[0.087,0.208]	[0.076,0.15]	[0.065,0.139]	[0.03,0.124]	[0.047,0.133]

Notes: This Table shows the estimates of the investment functions at different ages. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

positive and statistically different from zero. The significance of the income variable in the investment equation is important because income drives the identification of the investment coefficient in the production function. While, as we have discussed, the exclusion restriction that states that the impact of income on child development only works through parental investment might be questionable, at least our instrument is not a *weak* one.

2.5.2 The Production Function of Human Development

Having estimated the investment function, we can move on to the estimates we obtain for the production function for the three factors we consider. In what follows, we report the results for the Cobb-Doublas specification. We have explored richer specifications, such as a Translog and a CES specification, but, in most cases, we could not reject the restrictions implied by the Cobb Douglas. And even when there were deviations from the Cobb-Douglas specification, the rejections were marginally significant and had relatively small quantitative implications. We therefore concluded that the Cobb-Douglas specification constitutes a good approximation to our data.

This approach allows us to focus on three important issues. First, whether allowing for endogenous investment makes a difference in practice. Second, whether a Markov structure, where child outcomes depend on current development or a richer dynamic structure is necessary for the ages we consider. Third, we can study how the parameters of the production function change over time. These issues are particularly important to identify *windows of opportunities* for potential interventions.

For the three dimensions of child development (cognition, socio-emotional skills and health) we have two tables: one in which we consider only the current level of the relevant domain and one when we add an additional lag of the relevant domain. In all cases we also consider the effect of other domains on the relevant domain (for instance the effect of current socio-emotional skills and health on cognitive development). While we did explore the possibility of additional lags in these cross-effects, as we did not find any significant effects of this type, we do not report those results to avoid crowding the tables and making the estimates less precise.

In each of the tables below, we consider two panels, the left one in which the endogeneity of investment is ignored and the right one where we use the control function approach discussed above. This structure will allow us to consider both the importance of considering richer dynamics and the importance of considering endogenous investment. On this last issue, we should stress that we are limited by the nature of the data in terms of the instruments we can use. Effectivelly, our only alternative is to take the model we sketched in Section 2.2.4 literally and use parental income (which enters the budget constraint but not the production function) as an instrument. This choice can obviously be criticised, as income might be correlated with omitted inputs in the production function. As already discussed earlier, we notice that while in another data set from Colombia, Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020) use prices and exposure to conflict (which vary across localities) as exclusion restrictions to identify the impact of parental investment. Their results do not change much when they use only household resources, measured in a way similar to the variables we use, as an instrument.

Cognition at Different Ages

We start with a discussion of the results we obtain for the production function of cognitive development. The estimation results are reported in Tables 2.5 and 2.6. The left panel of the Tables contains the OLS estimates, which ignore the endogeneity of maternal investment, while the right-hand-side panel reports the estimates of the production function which include the control function derived from the investment function in Table 2.4. Table 2.5 contains the estimates of a specification that, for each age, allows only one lag of different dimensions of development in the production function for cognitive skills, while Table 2.6 allows for an additional lag of cognitive skills.[14] Finally, we notice that the two panels report the estimates obtained at each age between 2 and 7. The estimation of a production function for child development at this frequency is an important novelty of our paper and it is important as it helps to identify ages and periods in which investment might be particularly important, as we discuss in the next session. It also identifies the ages when the current state of development becomes particularly relevant for subsequent development.

We first notice that the dynamics of the process changes considerably with age, as can be deduced from the coefficient on lagged cognitive development. Up to age 4 included, the coefficient on lagged cognition are significantly different from zero, but the process is far from being very persistent, with coefficients around 0.3. This result holds both in the OLS and CF estimates and whether an additional lag is included or not. At age 5, however, the coefficient on lagged cognition increases considerably and become close to 1. At ages 6 and 7, the dynamics becomes more complex, as the coefficient on the second lag is also significantly different from zero and negative. The sum of the two coefficients is, from age 5 onward, close to 1.[15]

The coefficients on lagged socio-emotional skills are small in magnitude and only statistically significant at age 3 and 4. Instead, the coefficient on lagged health status is, at some ages, significant and important: lagged health seems to be important for cognitive development in particular at ages 2 and 4. At age 3, the coefficient on lagged health is also marginally significant, although considerably smaller in size than at the two adjacent ages. As we will see, these results turn out to be important for the dynamic targeting of investment and interventions.

Turning to the coefficients on parental investment, we notice that this variable appears significant and important only between ages 2 and 4. At this ages, we notice that the size of the coefficient at ages 3 and 4 doubles when we consider investment as endogenous, a result that is consistent with those in Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020), Attanasio, Meghir, & Nix (2020) and Attanasio, Meghir, Nix, & Salvati (2017). Perhaps surprisingly, the coefficient on parental investment at age 5 and older is not significantly different from zero. This result holds regardless of whether we treat investment as endogenous or not and regardless of the dynamic structure imposed on the production function. Moreover, the fact that the coefficient is not different from zero is not driven by an decrease in the precision of the estimates, but by a clear reduction of the point estimates.

Parental background does not seem to have a direct and significant role in the production function of cognitive development. We notice, however, that it does play an important role in parental investment, so that its effect on child development is mediated through investment, a result consistent with what reported by Attanasio,

[14] As mentioned above, we tried specifications with additional lags in the other two dimensions we consider (health and socio-emotional skills), but none of the coefficients was significantly different from zero. These results are available upon request.

[15] At age 2, we cannot have two lags of cognitive development for the obvious reason that we do not observe child cognitive development before birth.

Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020). Finally, we notice that the intervention that changed the nature of the care the children received in the nursery they were attending, has a positive and significant, although moderate, effect at ages 4 and 5. The point estimates of the effect at age 7 is of similar magnitude but, as the estimate is less precise, it is not statistically significant.

Socio-emotional Skills at Different Ages

We now move to the results obtained in estimating the production function for socio-emotional skills. We report the results in Tables 2.7 and 2.8 in a fashion similar to the results for the production function of cognitive skills, except for the fact that, due to the availability of appropriate measures, we only consider socio-emotional skills up to age 6. Before delving into a discussion of our estimates, it may be worth mentioning that the quality of the available measures of socio-emotional skills is not as high as that of other dimensions of development, as can be inferred from Table 2.2. For only two measures (one at age 1 and one at age 5) the signal to noise ration is above 0.5. As a consequence, our estimates are likely to be less informative and precise.

Starting with the dynamics of the process, we notice that the persistence of the process does not increase monotonically with age as with cognitive skills. In Table 2.7 the coefficient on lagged socio-emotional skills is around 0.3, with the exception of age 3, where is estimated (rather imprecisely) at 0.7. As for richer dynamics, Table 2.8 shows that the coefficient on the second lag is only significant at age 5 and, unlike for cognitive skills, is positive, although not very large. We notice that health has a positive impact on child socio-emotional skills at age 2 and age 5.

Moving on to the investment coefficients, we see that the control function is almost always significant, providing evidence for the endogeneity of investment. As the coefficient on the control function is negative, considering investment as endogenous increases (considerably) the size of its effect. In Table 2.7, for instance, at age 4 the point estimate of the investment coefficient goes from 0.135 to 0.506. Interestingly, the size of the coefficient on parental investment does not decline significantly with age, unlike with cognitive skills, where at age 5 and older it was effectively zero. In this case, at age 5 is still very high at 0.654. At age 6, it drop s to 0.354 but it is still significant.

We also notice that when considering parental investment as endogenous, parental cognition is not significant in the production function, a result that mirrors the one for cognitive skills.

Health at Different Ages

We now discuss the estimates of the production function of the health dimension of child development. Again we report two different tables, Tables 2.9 and 2.10, the first with just one lag of the three dimensions of development we are considering and the second with an additional lag for health. In each of the two Tables, the left panel consider parental investment as exogenous while on the right panel we use the control function approach to account for the possible endogeneity of investments. As for cognition, we estimate the function for ages 2 to 7.

Starting with the discussion of the dynamic properties, we notice that at some ages, both socio-emotional and cognitive skills appear marginally significant. However, the size of these coefficients is quite small, so that their effect is at best marginal. Lagged health status, however, is very important, signaling a strong persistence of the process. Already at age 2, the coefficient is larger than 0.7. At older ages, the coefficient reaches 1. At ages 5, 6 and 7, the additional lag attracts a negative and sizeable coefficient, which, for ages 6 and 7, is significantly different from zero. At these ages, the sum of the coefficients on the two lags of health is close to 1. The process, therefore, seems to have a complex dynamics.

Moving to the effect of parental investment, we notice that the only age at which it is productive is age 2, when it also seems to be the only age at which it should be considered as endogenous. As with other estimates of production functions, the control function takes a negative sign at age 2, so that its introduction increases the estimate of the coefficient on parental investment from 0.27 to 0.37. Past age 2, parental investment does not attract coefficients significantly different from zero.

We notice that parental cognition does not have a direct effect in the health production function, except at ages 6 and 7, when the intervention is marginally significant. For parental cognition, we remind the reader that it has an important effect on parental investment, which in turn is significant at age 2. We also notice that, given the high persistence of the process, an impact at an early age, can be traced in subsequent outcomes. Finally, we notice that the $aeioTU$ intervention has a small and positive effect on child health (marginally significant at age 6 and 7).

Table 2.5: Production Functions for Cognitive Skills

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.155	-0.135	0.44	0.552	0.74	-0.29	0.027	-0.176	0.481	0.548	0.739
	0.16	0.11	0.03	0.054	0.043	0.05	0.154	0.114	0.044	0.07	0.05	0.05
	[-0.443,0.079]	[0.048,0.402]	[-0.191,-0.098]	[0.341,0.517]	[0.486,0.624]	[0.651,0.811]	[-0.52,-0.037]	[-0.097,0.299]	[-0.278,-0.145]	[0.37,0.588]	[0.456,0.626]	[0.649,0.813]
Control function	-	-	-	-	-	-	-0.22	-0.268	-0.132	0.133	-0.028	-0.061
							0.145	0.102	0.098	0.113	0.161	0.12
							[-0.423,0.036]	[-0.43,-0.127]	[-0.366,-0.053]	[-0.017,0.333]	[-0.272,0.193]	[-0.223,0.158]
AeioTu	-0.13	0.034	0.03	0.068	0.025	0.05	-0.11	0.046	0.03	0.068	0.026	0.056
	0.081	0.056	0.013	0.018	0.045	0.042	0.078	0.051	0.013	0.018	0.046	0.045
	[-0.217,0.038]	[-0.048,0.125]	[0.007,0.047]	[0.026,0.09]	[-0.048,0.099]	[-0.004,0.122]	[-0.208,0.046]	[-0.036,0.129]	[0.007,0.046]	[0.026,0.088]	[-0.051,0.093]	[-0.005,0.134]
Cognitive t	0.3	0.501	0.436	1.021	1.138	0.908	0.3	0.494	0.423	1.051	1.133	0.898
	0.067	0.081	0.039	0.085	0.059	0.048	0.067	0.082	0.042	0.1	0.072	0.054
	[0.229,0.442]	[0.449,0.692]	[0.372,0.498]	[0.879,1.141]	[1.01,1.205]	[0.805,0.954]	[0.221,0.442]	[0.421,0.678]	[0.349,0.487]	[0.901,1.207]	[0.982,1.22]	[0.79,0.969]
Socio-emotional t	0.08	0.158	0.058	-0.06	0.007	0.058	0.08	0.121	0.046	-0.043	0.002	0.053
	0.103	0.075	0.023	0.031	0.036	0.043	0.101	0.074	0.028	0.035	0.049	0.044
	[-0.007,0.335]	[0.018,0.263]	[0.008,0.085]	[-0.089,0.014]	[-0.074,0.044]	[-0.001,0.133]	[-0.014,0.326]	[-0.022,0.215]	[-0.017,0.071]	[-0.071,0.038]	[-0.1,0.063]	[-0.012,0.139]
Health t	0.39	0.238	0.293	0.006	-0.065	-0.062	0.35	0.161	0.267	0.024	-0.066	-0.066
	0.12	0.103	0.048	0.063	0.044	0.034	0.116	0.097	0.055	0.063	0.045	0.035
	[0.134,0.512]	[0.051,0.366]	[0.193,0.362]	[-0.087,0.108]	[-0.119,0.021]	[-0.095,0.011]	[0.108,0.478]	[-0.003,0.319]	[0.156,0.327]	[-0.067,0.137]	[-0.123,0.02]	[-0.103,0.009]
Parental cognition	0.02	0.013	0.102	-0.011	-0.031	0.101	0	-0.008	0.061	0.025	-0.04	0.079
	0.089	0.054	0.031	0.036	0.045	0.045	0.089	0.061	0.044	0.046	0.068	0.064
	[-0.017,0.233]	[-0.011,0.142]	[0.058,0.159]	[-0.051,0.065]	[-0.085,0.064]	[-0.008,0.142]	[-0.047,0.223]	[-0.103,0.063]	[-0.027,0.118]	[-0.023,0.121]	[-0.138,0.086]	[-0.045,0.16]
Investments	0.2	0.091	0.111	0.043	-0.05	-0.005	0.27	0.231	0.203	-0.057	-0.03	0.036
	0.085	0.068	0.041	0.042	0.053	0.045	0.12	0.103	0.092	0.103	0.145	0.097
	[-0.017,0.28]	[-0.035,0.184]	[0.061,0.191]	[-0.06,0.081]	[-0.129,0.045]	[-0.032,0.106]	[-0.032,0.376]	[0.082,0.387]	[0.134,0.418]	[-0.273,0.04]	[-0.234,0.232]	[-0.113,0.19]

Notes: This table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.11 for coefficients on these covariates).

Table 2.6: Production Functions for Cognitive Skills with Two Self-productivity Lags

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.202	-0.056	0.44	0.422	0.483	-0.29	0.097	-0.098	0.484	0.389	0.478
	0.16	0.102	0.043	0.047	0.073	0.109	0.154	0.101	0.047	0.062	0.083	0.11
	[-0.443,0.079]	[0.1,0.426]	[-0.121,0.012]	[0.353,0.505]	[0.342,0.575]	[0.527,0.876]	[-0.52,-0.037]	[-0.016,0.326]	[-0.199,-0.044]	[0.386,0.574]	[0.298,0.565]	[0.516,0.871]
Control function	-	-	-	-	-	-	-0.22	-0.303	-0.147	0.136	-0.135	-0.077
							0.145	0.107	0.093	0.112	0.157	0.122
							[-0.423,0.036]	[-0.459,-0.124]	[-0.361,-0.066]	[-0.011,0.327]	[-0.395,0.12]	[-0.235,0.145]
AeioTu	-0.13	0.034	0.056	0.068	0.006	0.046	-0.11	0.048	0.057	0.068	0.01	0.054
	0.081	0.057	0.016	0.018	0.046	0.043	0.078	0.051	0.016	0.018	0.047	0.045
	[-0.217,0.038]	[-0.049,0.127]	[0.028,0.08]	[0.025,0.089]	[-0.061,0.085]	[-0.009,0.125]	[-0.208,0.046]	[-0.044,0.127]	[0.028,0.081]	[0.026,0.087]	[-0.055,0.087]	[-0.013,0.132]
Cognitive t	0.3	0.469	0.319	1.022	1.281	1.321	0.3	0.435	0.3	1.044	1.271	1.315
	0.067	0.124	0.069	0.124	0.092	0.178	0.067	0.125	0.071	0.133	0.098	0.181
	[0.229,0.442]	[0.379,0.741]	[0.204,0.427]	[0.808,1.215]	[1.072,1.365]	[0.69,1.211]	[0.221,0.442]	[0.348,0.707]	[0.161,0.403]	[0.834,1.27]	[1.046,1.362]	[0.683,1.213]
Cognitive t-1	-	0.046	0.155	0	-0.248	-0.507	-	0.084	0.162	0.011	-0.275	-0.514
		0.076	0.066	0.073	0.115	0.199		0.075	0.065	0.073	0.117	0.202
		[-0.094,0.134]	[0.067,0.276]	[-0.12,0.101]	[-0.365,0.019]	[-0.377,0.205]		[-0.077,0.147]	[0.071,0.279]	[-0.111,0.121]	[-0.402,-0.014]	[-0.39,0.202]
Socio-emotional t	0.08	0.161	0.049	-0.06	0.053	0.082	0.08	0.123	0.035	-0.043	0.034	0.076
	0.103	0.077	0.023	0.031	0.044	0.042	0.101	0.075	0.027	0.035	0.052	0.043
	[-0.007,0.335]	[0.02,0.264]	[-0.002,0.074]	[-0.088,0.014]	[-0.061,0.085]	[0.006,0.143]	[-0.014,0.326]	[-0.017,0.216]	[-0.027,0.06]	[-0.072,0.039]	[-0.096,0.067]	[-0.012,0.14]
Health t	0.39	0.213	0.269	0.006	-0.03	-0.037	0.35	0.107	0.238	0.02	-0.032	-0.041
	0.12	0.1	0.047	0.051	0.045	0.035	0.116	0.095	0.054	0.053	0.046	0.036
	[0.134,0.512]	[0.034,0.355]	[0.178,0.339]	[-0.066,0.111]	[-0.099,0.044]	[-0.095,0.015]	[0.108,0.478]	[-0.017,0.271]	[0.132,0.309]	[-0.055,0.124]	[-0.105,0.044]	[-0.1,0.018]
Parental cognition	0.02	0.016	0.092	-0.011	0.006	0.12	0	-0.006	0.045	0.026	-0.035	0.092
	0.089	0.059	0.029	0.036	0.048	0.048	0.089	0.062	0.041	0.046	0.067	0.067
	[-0.017,0.233]	[-0.01,0.161]	[0.046,0.144]	[-0.054,0.064]	[-0.065,0.096]	[-0.009,0.149]	[-0.047,0.223]	[-0.092,0.087]	[-0.038,0.094]	[-0.023,0.117]	[-0.14,0.085]	[-0.048,0.171]
Investments	0.2	0.095	0.117	0.043	-0.06	0.02	0.27	0.258	0.221	-0.058	0.036	0.072
	0.085	0.068	0.038	0.044	0.053	0.046	0.12	0.106	0.086	0.103	0.143	0.099
	[-0.017,0.28]	[-0.038,0.176]	[0.066,0.194]	[-0.068,0.083]	[-0.131,0.049]	[-0.036,0.118]	[-0.032,0.376]	[0.088,0.401]	[0.147,0.416]	[-0.275,0.042]	[-0.165,0.297]	[-0.096,0.221]

Notes: This Table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.12 for coefficients on these covariates).

Table 2.7: Production Functions for Socio-emotional Skills

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.104	0.217	0.312	0.286	0.11	-0.12	0.054	0.09	0.236
	0.303	0.408	0.07	0.064	0.079	0.326	0.409	0.104	0.108	0.089
	[-0.143,0.812]	[-0.596,0.706]	[0.1,0.327]	[0.222,0.451]	[0.158,0.407]	[-0.483,0.597]	[-0.893,0.355]	[-0.141,0.175]	[-0.102,0.251]	[0.102,0.377]
Control function	-	-	-	-	-	-0.62	-0.471	-0.529	-0.728	-0.356
						0.279	0.263	0.237	0.273	0.348
						[-0.979,-0.097]	[-1.045,-0.244]	[-0.955,-0.241]	[-1.288,-0.368]	[-0.947,0.169]
AeioTu	-0.12	0.13	-0.007	-0.115	-0.001	-0.08	0.151	-0.006	-0.117	0.017
	0.181	0.162	0.038	0.031	0.1	0.173	0.156	0.037	0.031	0.106
	[-0.375,0.234]	[-0.305,0.219]	[-0.072,0.05]	[-0.175,-0.072]	[-0.195,0.116]	[-0.343,0.244]	[-0.25,0.218]	[-0.07,0.046]	[-0.172,-0.071]	[-0.201,0.154]
Cognitive t	-0.15	-0.007	0.216	0.099	0.217	-0.16	-0.019	0.167	-0.064	0.151
	0.131	0.206	0.076	0.078	0.088	0.133	0.206	0.088	0.129	0.118
	[-0.369,0.05]	[-0.552,0.138]	[0.062,0.319]	[-0.019,0.224]	[0.072,0.359]	[-0.391,0.035]	[-0.577,0.121]	[-0.019,0.282]	[-0.321,0.083]	[-0.044,0.33]
Socio-emotional t	0.28	0.77	0.41	0.436	0.378	0.26	0.707	0.361	0.345	0.316
	0.159	0.216	0.077	0.067	0.092	0.148	0.229	0.089	0.075	0.109
	[0.118,0.609]	[0.304,0.983]	[0.332,0.568]	[0.332,0.543]	[0.242,0.548]	[0.106,0.538]	[0.208,0.912]	[0.258,0.538]	[0.211,0.454]	[0.148,0.495]
Health t	0.42	0.09	0.156	0.242	0.103	0.31	-0.044	0.05	0.144	0.09
	0.176	0.198	0.082	0.068	0.073	0.184	0.21	0.087	0.086	0.077
	[0.122,0.685]	[-0.106,0.556]	[0.051,0.315]	[0.16,0.38]	[-0.017,0.219]	[0.012,0.605]	[-0.287,0.425]	[-0.077,0.208]	[0.022,0.296]	[-0.069,0.201]
Parental cognition	0.04	-0.021	0.083	0.115	0.206	-0.01	-0.058	-0.084	-0.079	0.088
	0.111	0.129	0.068	0.056	0.089	0.152	0.146	0.095	0.094	0.149
	[-0.071,0.291]	[-0.054,0.37]	[-0.068,0.15]	[0.03,0.217]	[0.054,0.336]	[-0.323,0.145]	[-0.309,0.103]	[-0.309,-0.007]	[-0.231,0.06]	[-0.172,0.287]
Investments	0.41	0.167	0.135	0.108	0.095	0.6	0.415	0.506	0.654	0.354
	0.163	0.137	0.068	0.083	0.099	0.232	0.246	0.203	0.243	0.286
	[0.104,0.593]	[-0.052,0.386]	[0.043,0.264]	[-0.054,0.202]	[-0.049,0.257]	[0.259,0.944]	[0.223,0.989]	[0.256,0.92]	[0.368,1.136]	[-0.049,0.869]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.13 for coefficients on these covariates).

Table 2.8: Production Functions for Socio-emotional Skills with Two Self-productivity Lags

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.03	0.175	0.316	0.31	0.11	-0.079	0.06	0.107	0.272
	0.303	0.535	0.076	0.065	0.078	0.326	0.508	0.102	0.106	0.088
	[-0.143,0.812]	[-0.997,0.783]	[0.034,0.271]	[0.226,0.438]	[0.195,0.452]	[-0.483,0.597]	[-1.126,0.534]	[-0.133,0.177]	[-0.053,0.268]	[0.159,0.424]
Control function	-	-	-	-	-	-0.62	-0.258	-0.447	-0.683	-0.242
						0.279	0.243	0.25	0.268	0.347
						[-0.979,-0.097]	[-0.747,0.012]	[-0.935,-0.127]	[-1.228,-0.303]	[-0.749,0.369]
AeioTu	-0.12	0.147	0.035	-0.1	0.01	-0.08	0.158	0.012	-0.105	0.021
	0.181	0.15	0.038	0.033	0.097	0.173	0.148	0.039	0.034	0.102
	[-0.375,0.234]	[-0.222,0.24]	[-0.034,0.093]	[-0.158,-0.05]	[-0.18,0.125]	[-0.343,0.244]	[-0.217,0.271]	[-0.058,0.062]	[-0.16,-0.059]	[-0.191,0.142]
Cognitive t	-0.15	-0.041	0.212	0.093	0.17	-0.16	-0.04	0.173	-0.059	0.133
	0.131	0.272	0.077	0.081	0.094	0.133	0.266	0.088	0.124	0.116
	[-0.369,0.05]	[-0.721,0.182]	[0.051,0.31]	[-0.032,0.23]	[-0.003,0.312]	[-0.391,0.035]	[-0.716,0.138]	[-0.018,0.282]	[-0.286,0.1]	[-0.064,0.313]
Socio-emotional t	0.28	0.728	0.39	0.338	0.314	0.26	0.702	0.36	0.275	0.282
	0.159	0.211	0.085	0.076	0.103	0.148	0.224	0.09	0.078	0.111
	[0.118,0.609]	[0.245,0.918]	[0.293,0.557]	[0.193,0.448]	[0.149,0.484]	[0.106,0.538]	[0.173,0.887]	[0.248,0.534]	[0.116,0.369]	[0.099,0.46]
Socio-emotional t-1	-	0.087	0.171	0.114	0.164	-	0.066	0.075	0.088	0.139
		0.246	0.093	0.079	0.119		0.246	0.096	0.078	0.125
		[-0.265,0.61]	[0.034,0.3]	[0.054,0.283]	[0.013,0.393]		[-0.243,0.588]	[-0.081,0.215]	[0.015,0.242]	[-0.024,0.383]
Health t	0.42	0.078	0.118	0.263	0.093	0.31	0.005	0.05	0.166	0.086
	0.176	0.203	0.083	0.071	0.072	0.184	0.21	0.085	0.087	0.074
	[0.122,0.685]	[-0.121,0.555]	[0.003,0.268]	[0.162,0.397]	[-0.025,0.205]	[0.012,0.605]	[-0.246,0.468]	[-0.073,0.203]	[0.031,0.318]	[-0.043,0.196]
Parental cognition	0.04	-0.024	0.065	0.111	0.178	-0.01	-0.044	-0.066	-0.07	0.102
	0.111	0.112	0.066	0.055	0.088	0.152	0.122	0.097	0.09	0.144
	[-0.071,0.291]	[-0.097,0.249]	[-0.084,0.13]	[0.016,0.203]	[0.007,0.302]	[-0.323,0.145]	[-0.256,0.101]	[-0.28,0.018]	[-0.227,0.062]	[-0.148,0.312]
Investments	0.41	0.173	0.044	0.082	0.081	0.6	0.311	0.409	0.6	0.259
	0.163	0.15	0.071	0.089	0.095	0.232	0.22	0.22	0.247	0.28
	[0.104,0.593]	[-0.029,0.45]	[-0.042,0.199]	[-0.101,0.171]	[-0.067,0.245]	[0.259,0.944]	[0.087,0.843]	[0.204,0.895]	[0.222,1.072]	[-0.189,0.697]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.14 for coefficients on these covariates).

Table 2.9: Production Functions for Health

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.382	0.347	0.38	0.331	0.496	-0.06	0.402	0.368	0.353	0.35	0.497
	0.145	0.062	0.017	0.021	0.026	0.026	0.135	0.071	0.021	0.025	0.033	0.027
	[-0.14,0.319]	[0.252,0.449]	[0.323,0.379]	[0.338,0.407]	[0.305,0.391]	[0.455,0.538]	[-0.288,0.166]	[0.235,0.47]	[0.341,0.409]	[0.307,0.391]	[0.323,0.425]	[0.455,0.539]
Control function	-	-	-	-	-	-	-0.33	0.042	0.067	-0.09	0.132	0.033
							0.116	0.061	0.037	0.046	0.137	0.098
							[-0.473,-0.119]	[-0.102,0.106]	[0.015,0.128]	[-0.13,0.011]	[-0.012,0.433]	[-0.061,0.236]
AeioTu	-0.03	0.004	0.005	0.005	0.069	0.105	0	0.002	0.005	0.005	0.062	0.102
	0.061	0.042	0.005	0.007	0.032	0.028	0.05	0.041	0.005	0.007	0.032	0.03
	[-0.104,0.086]	[-0.069,0.059]	[-0.002,0.016]	[-0.009,0.014]	[0.004,0.105]	[0.05,0.137]	[-0.076,0.076]	[-0.068,0.06]	[-0.002,0.016]	[-0.01,0.014]	[-0.005,0.097]	[0.04,0.133]
Cognitive t	-0.02	0.061	0.002	-0.13	-0.1	-0.122	-0.03	0.062	0.008	-0.151	-0.075	-0.117
	0.063	0.1	0.022	0.033	0.038	0.033	0.062	0.098	0.022	0.036	0.049	0.036
	[-0.108,0.096]	[-0.043,0.268]	[-0.037,0.037]	[-0.191,-0.086]	[-0.152,-0.031]	[-0.177,-0.072]	[-0.108,0.093]	[-0.043,0.26]	[-0.03,0.044]	[-0.214,-0.093]	[-0.126,0.026]	[-0.171,-0.05]
Socio-emotional t	-0.01	0.069	0.026	0.042	-0.037	-0.035	-0.02	0.075	0.032	0.031	-0.014	-0.032
	0.098	0.052	0.011	0.014	0.036	0.027	0.087	0.05	0.014	0.016	0.049	0.028
	[-0.1,0.225]	[0.003,0.18]	[0.006,0.041]	[0.009,0.055]	[-0.088,0.028]	[-0.085,0.005]	[-0.1,0.187]	[0.006,0.169]	[0.012,0.057]	[-0.005,0.052]	[-0.07,0.079]	[-0.078,0.021]
Health t	0.77	0.892	0.95	1.107	1.196	1.237	0.71	0.904	0.963	1.095	1.202	1.239
	0.107	0.155	0.028	0.034	0.043	0.034	0.097	0.162	0.03	0.034	0.045	0.036
	[0.533,0.872]	[0.543,1.036]	[0.914,1.009]	[1.051,1.164]	[1.103,1.249]	[1.173,1.291]	[0.495,0.822]	[0.522,1.056]	[0.925,1.022]	[1.047,1.159]	[1.11,1.263]	[1.181,1.301]
Parental cognition	0	-0.006	-0.023	0.029	-0.017	-0.071	-0.03	-0.002	-0.002	0.005	0.027	-0.06
	0.071	0.029	0.015	0.02	0.035	0.032	0.058	0.031	0.018	0.023	0.063	0.049
	[-0.017,0.189]	[-0.026,0.062]	[-0.047,0.003]	[-0.013,0.052]	[-0.075,0.031]	[-0.123,-0.017]	[-0.062,0.095]	[-0.028,0.062]	[-0.03,0.03]	[-0.04,0.036]	[-0.049,0.16]	[-0.118,0.034]
Investments	0.27	-0.017	0.045	-0.048	-0.043	-0.009	0.37	-0.039	-0.002	0.02	-0.139	-0.031
	0.073	0.045	0.019	0.022	0.04	0.035	0.083	0.064	0.034	0.041	0.13	0.082
	[0.061,0.301]	[-0.061,0.082]	[0.01,0.069]	[-0.059,0.014]	[-0.102,0.023]	[-0.053,0.061]	[0.18,0.443]	[-0.117,0.099]	[-0.063,0.046]	[-0.038,0.101]	[-0.401,0.006]	[-0.196,0.061]

Notes: This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.15 for coefficients on these covariates).

Table 2.10: Production Functions for Health with Two Self-productivity Lags

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.394	0.298	0.256	0.076	0.376	-0.06	0.407	0.318	0.222	0.08	0.371
	0.145	0.065	0.026	0.072	0.095	0.059	0.135	0.068	0.028	0.079	0.097	0.062
	[-0.14,0.319]	[0.252,0.472]	[0.261,0.347]	[0.298,0.53]	[-0.022,0.28]	[0.306,0.486]	[-0.288,0.166]	[0.249,0.479]	[0.282,0.369]	[0.266,0.517]	[-0.003,0.3]	[0.301,0.498]
Control function	-	-	-	-	-	-	-0.33	0.034	0.06	-0.1	0.204	-0.032
							0.116	0.07	0.038	0.048	0.12	0.099
							[-0.473,-0.119]	[-0.112,0.101]	[0.002,0.12]	[-0.13,0.018]	[0.027,0.411]	[-0.125,0.184]
AeioTu	-0.03	0.004	0.009	0.001	0.028	0.074	0	0.003	0.009	0.001	0.014	0.076
	0.061	0.042	0.006	0.007	0.03	0.029	0.05	0.042	0.006	0.007	0.03	0.03
	[-0.104,0.086]	[-0.068,0.058]	[0.002,0.02]	[-0.008,0.014]	[-0.026,0.068]	[0.029,0.121]	[-0.076,0.076]	[-0.069,0.059]	[0.002,0.021]	[-0.008,0.014]	[-0.034,0.061]	[0.027,0.119]
Cognitive t	-0.02	0.05	0.012	-0.092	-0.032	-0.092	-0.03	0.054	0.017	-0.113	0.012	-0.096
	0.063	0.084	0.023	0.042	0.041	0.033	0.062	0.082	0.023	0.043	0.048	0.036
	[-0.108,0.096]	[-0.069,0.212]	[-0.031,0.047]	[-0.22,-0.077]	[-0.091,0.042]	[-0.157,-0.045]	[-0.108,0.093]	[-0.062,0.203]	[-0.025,0.054]	[-0.24,-0.094]	[-0.054,0.095]	[-0.158,-0.044]
Socio-emotional t	-0.01	0.067	0.025	0.038	-0.007	-0.01	-0.02	0.072	0.031	0.025	0.031	-0.012
	0.098	0.051	0.012	0.014	0.039	0.029	0.087	0.049	0.014	0.017	0.045	0.028
	[-0.1,0.225]	[0.002,0.168]	[0.005,0.041]	[0.008,0.056]	[-0.058,0.063]	[-0.068,0.032]	[-0.1,0.187]	[0.005,0.153]	[0.008,0.053]	[-0.004,0.054]	[-0.03,0.113]	[-0.063,0.028]
Health t	0.77	0.857	1.051	1.472	1.752	1.546	0.71	0.876	1.06	1.47	1.815	1.554
	0.107	0.233	0.066	0.215	0.208	0.152	0.097	0.248	0.067	0.217	0.206	0.16
	[0.533,0.872]	[0.381,1.119]	[0.955,1.175]	[0.639,1.327]	[1.278,1.97]	[1.26,1.744]	[0.495,0.822]	[0.334,1.132]	[0.965,1.186]	[0.654,1.335]	[1.323,2.006]	[1.226,1.743]
Health t-1	-	0.041	-0.117	-0.385	-0.698	-0.396	-	0.03	-0.113	-0.397	-0.766	-0.409
		0.144	0.06	0.223	0.252	0.176		0.149	0.061	0.228	0.249	0.191
		[-0.178,0.272]	[-0.214,-0.021]	[-0.219,0.491]	[-0.976,-0.155]	[-0.626,-0.066]		[-0.192,0.286]	[-0.208,-0.02]	[-0.229,0.488]	[-1.005,-0.193]	[-0.635,-0.018]
Parental cognition	0	-0.006	-0.019	0.034	0.012	-0.063	-0.03	-0.003	0	0.007	0.082	-0.074
	0.071	0.033	0.016	0.02	0.036	0.033	0.058	0.032	0.019	0.023	0.056	0.045
	[-0.017,0.189]	[-0.026,0.079]	[-0.041,0.011]	[-0.015,0.05]	[-0.053,0.056]	[-0.116,-0.008]	[-0.062,0.095]	[-0.028,0.063]	[-0.027,0.034]	[-0.038,0.036]	[-0.015,0.172]	[-0.119,0.014]
Investments	0.27	-0.009	0.047	-0.068	-0.027	0.015	0.37	-0.029	0.005	0.007	-0.174	0.036
	0.073	0.053	0.019	0.028	0.037	0.038	0.083	0.081	0.034	0.042	0.109	0.081
	[0.061,0.301]	[-0.069,0.095]	[0.01,0.073]	[-0.065,0.025]	[-0.089,0.028]	[-0.039,0.076]	[0.18,0.443]	[-0.119,0.141]	[-0.059,0.053]	[-0.035,0.107]	[-0.382,-0.037]	[-0.122,0.122]

Notes: This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table 2.16 for coefficients on these covariates).

2.6 *Windows of Opportunity*

In this section, we use the model to better characterise the process of human development in its various dimensions, focusing on aspects that might be relevant for the design of policy. In particular, we want to establish whether the nature of the process of development points to the existence of *windows of opportunities*, that is periods in which interventions (or investment) could be particular effective and their effect sustainable. To do so, we perform some simulations where we force some changes to the inputs in the production function of human development and then study the long term effects of such changes by looking at the *impulse response function* of these changes for the variables we consider. This exercise is particularly useful because it takes into account all dynamic interactions between the various components of development as well as the effect of parental investment. Before performing the simulations, however, we present some evidence on how well the model fits the data and some graphical representation of the marginal productivity of parental investment at different ages.

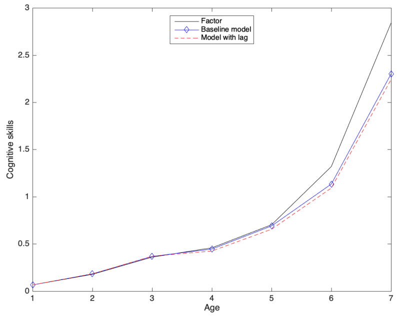
2.6.1 Model Fit

We start to use our estimates of the production functions for the three dimensions of human development we consider at the various ages, checking how well the model fits the data on which it was estimated. In Figure 2.3, we plot the evolution of cognitive skills, socio-emotional skills and health over time as predicted by our model estimates, as reported in Tables 2.6, 2.8 and 2.9, as well as the average factor corresponding to each dimension considerd, implied by the available measures and the estimated measurement system. The figures are constructed assigning the median initial conditions in our sample and holding investments fixed at their median value in each period. In each graph we plot both the estimates from the baseline model and the estimates from the model with self-productivity lags.

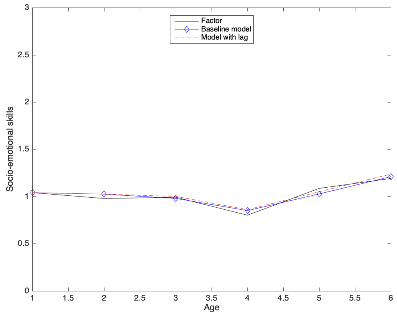
The figures show that, perhaps not surprisingly, our model fit is remarkably good. For the cognitive factor, however, both the baseline model and the one with an additional lag in cognition, underpredict child development at age 6 and, more severely, at age 7. For socio-emotional skills and health the fit of both versions of the model is remarkably good.

2.6.2 Marginal Product of Investment

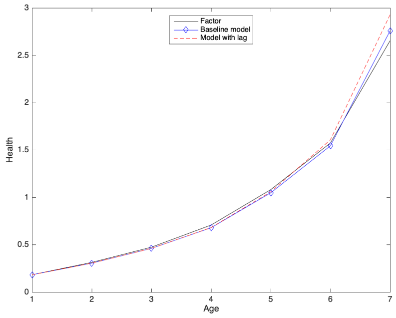
In Figure 2.4, we plot the marginal product of investment (MPI) by baseline skills level, computed using the estimates from the production functions for cognitive, socio-emotional skills and health, as implied by our model estimates. As before, the estimates used here are those reported in Tables 2.6, 2.8 and 2.9. The first three graphs plot the MPI for cognitive skills, the following three for socio-emotional skills, and the last three for health. Within each row, each column shows how the marginal product



(a) Cognitive skills



(b) Socio-emotional skills



(c) Health

Figure 2.3: Evolution of Skills Over Time (Model Estimates)

Notes: The figure plots the evolution of cognitive skills (panel a), socio-emotional skills (panel b), and health (panel c) over time as predicted by our model estimates reported in Tables 2.6, 2.8 and 2.9, as well as the average factor corresponding to each dimension considered. The figures are constructed assigning the median initial conditions in our sample and holding investments fixed at their median value in each period. In each graph we plot both the estimates from the baseline model (solid blue line) and the estimates from the model with self-productivity lags (dashed red line).

varies by baseline level of cognitive skills (column 1), socio-emotional skills (column 2) and health (column 3). In each graph, we plot the MPI at different ages. The figures are constructed evaluating the marginal product at different levels of baseline skills, while holding all other inputs at their median value in the sample. We note that the scale is different in each graph.

First, we note the complementarity of investment in each of the three dimensions with the relevant level of development (the graphs on the diagonal). These are evident in particular for cognition and socio-emotional skills, while for health it is only relevant for the production function at age 2. For cognition the complementarity (the slope of the marginal productivity lines) becomes more evident at age 3. Moreover, investment seems to be complementary also with health and, to an extent, with socio-emotional skills. These results point to the fact that investments in children with higher baseline skill levels are more productive. This highlights the importance of improving child development in the very early years in order for subsequent investments to be productive.

In the case of socio-emotional skills, there are no strong complementaries of investment with lagged cognitive development. However we do find evidence of complementarity between lagged health and parental investments in the production for socio-emotional skills at age 2. Finally, in the case of the health production function we find no complementaries between lagged cognition or socio-emotional skills and parental investments.

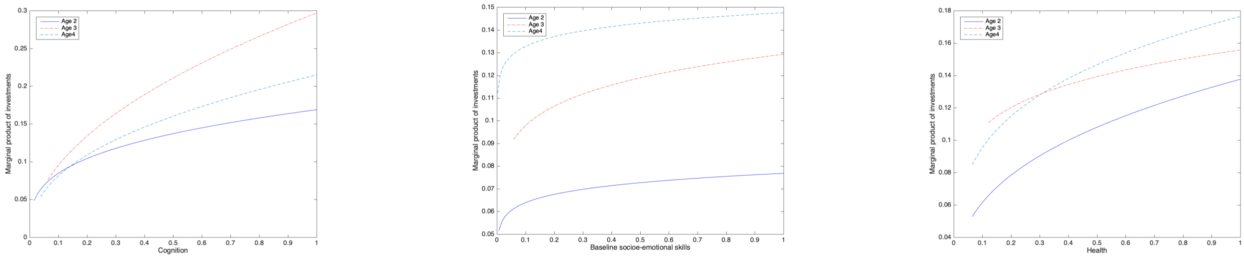
2.6.3 Impulse Response Functions

In Figure 2.5, we plot the impulse response functions (IRFs) for cognitive skills following an investment innovation, as implied by the estimated production functions. The IRFs are dynamic and take into account the effects that, over time, a given change in investment has on all dimensions of child development we consider. In each plot, we increase investments by one standard deviation at a particular age (panels a, b, c and d). In panel e, investment is increased at ages 2, 3 and 4, while in panel f, it is increased at ages 3 and 4. We compare the evolution of cognitive skills to what would have occurred if investments were fixed at their median level observed in the data. We assume that at baseline all other inputs are fixed at their median level. Each graph reports both the evolution of skills predicted by the production function that allows for only one lag of the various dimensions of child development and that predicted by the model with an additional self-productivity lag.

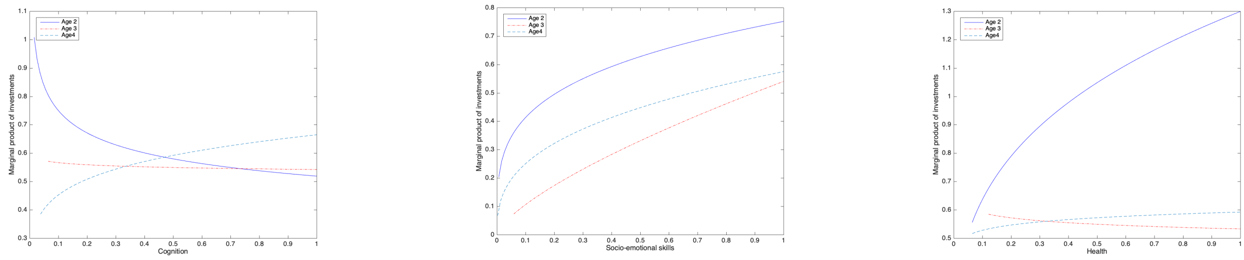
A striking difference is that observed between panels a and b. In the former, when investment increases at age 2 only, it gives an impact of about 12% of a standard deviation on cognitive skills. Such effect is sustained through age 7, although the model with a richer dynamics implies a slight decline in particular at ages 6 and 7. The picture is very different in the increase happens at age 3: while the magnitude of the effect is very similar, this vanishes almost completely in age 4. The main reason behind this differences is the different dynamics of the process at different ages. We notice that investment at age 2 has a considerable effect on the health dimension at age 2. This effect is then reflected on cognition at age 3 and all subsequent ages. This result highlights that early investments can achieve large impacts both by *directly* effecting cognitive skills, and *indirectly* affecting cognition thought the existence of complementaries with early health.

We also notice that investment at age 4 has a persistent effect, while the effect of parental investment at age 5 is absent. Finally, the comparison between panel e and f and the previous four panel makes clear not only the importance of starting investing early, but the importance of sustaining investments for several periods. The cumulative effect in panel e is much larger than that in panel f and that of short lived increases in investment considered in previous panels.

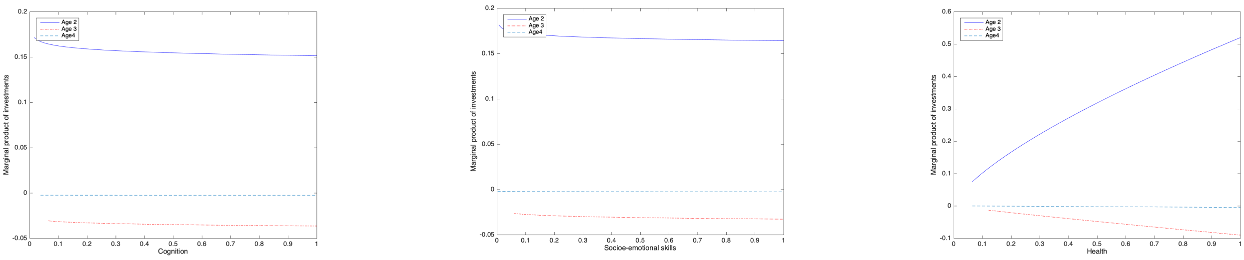
While we do not report confidence interval for these exercises, they show the importance of considering the process of child development as a whole in its various dimensions. The dynamic of this process can be complex and this has important implications for the identification of *windows of opportunities* to target interventions



(a) Cognitive skills



(b) Socio-emotional skills



(c) Health

Figure 2.4: Marginal Product of Investment

Notes: This figure plots marginal product of investment (MPI) by baseline skills level, computed using the estimates from the production functions for cognitive skills, socio-emotional skills and health, as implied by our model estimates. The top three panels plot the MPI for cognitive skills, the middle three for socio-emotional skills and the bottom three for health. Within each row, each column shows how the marginal product varies by baseline level of cognitive skills (column 1), socio-emotional skills (column 2) and health (column 3). In each graph, we plot the MPI at age 2 (solid blue), age 3 (dashed red) and age 4 (dashed blue). The figures are constructed evaluating the marginal product at different levels of baseline skills, while holding all other inputs at their median value in the sample. We note that the scale is different in each graph.

and policies to improve the chances of disadvantaged children.

In Figure 2.6, we repeat for socio-emotional skills the exercise performed for cognitive skills. Again, we increase investment exogenously by one standard deviation at different ages, following the same scheme we used for cognitive skills. We then compare the evolution of socio-emotional skills to what would have occurred if investments were fixed at their median level. Again, at baseline, all other inputs are fixed at their median level. Each graph reports both the evolution of skills predicted by our baseline model and that predicted by the model with an additional self-productivity lag.

While we do not find the dramatic difference in dynamics we observe for cognitive skills, again the impact of investment is much more pronounced at age 2 than at subsequent ages. The effects of the two models (with or without additional lags) are very similar. When looking at the effects of changing the investment at ages 2, 3 and 4, rather than only at ages 3 and 4, it is clear that the cumulate impact of the former is higher than that of the latter.

An impressive feature of the impacts on the accumulation of socio-emotional skills is the effect of investment at age 5, which is quite high and compares with a zero impact on cognitive skills at the same age. While we do not report the IRF for an increase at ages 2 to 5, it is clear from this picture and the coefficients in the production function, that the effect of such an intervention would be very high.

Finally, in Figure 2.7, we plot the impulse response functions (IRFs) for health following the same investment innovations considered for cognition and socio-emotional skills. As before, we assume that at baseline all other inputs are fixed at their median level and we report the effects implied by the two versions of the model we consider.

We notice that our estimates imply a persistent effect of investment on the health dimension of child development if done at age 2 but not for subsequent ages. Once again, we observe some differences in the long run effect of the two models in panel a but not in the other panels. As we discussed above, while not particularly surprising, given the coefficients of the health production function, these results, and in particular the effect of investment on health age age 2, are important to better understand the effect of investment on other dimensions of child development, and underscore the importance of preventing ill health from an early stage to improve child cognitive outcomes in the long run.

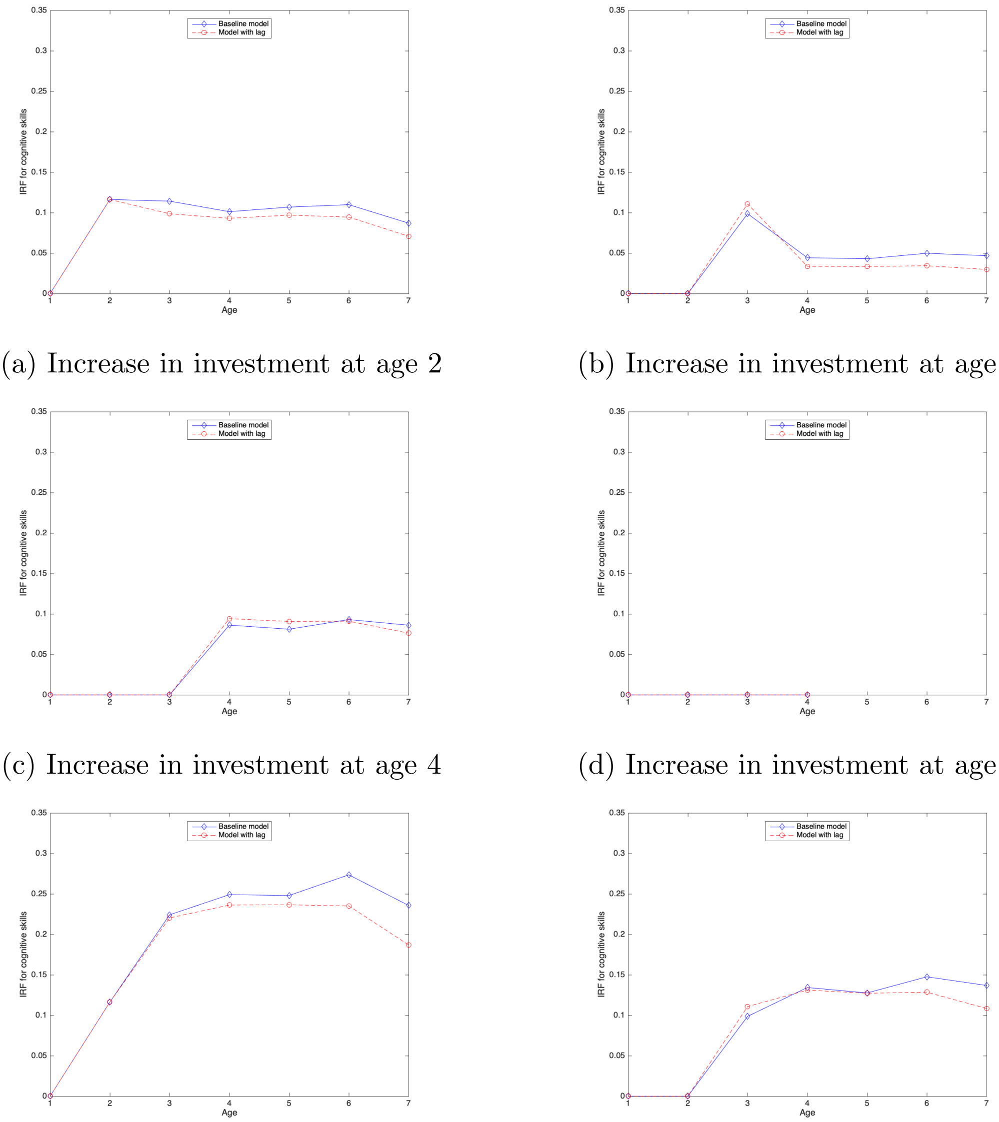
2.7 Conclusions

In this paper, we study the process of human development from age 1 to 7, in three important dimensions: cognition, socio-emotional skills and health. We use a rich and unique data set from Colombia, which contains high frequency information for a large sample of children from vulnerable families. We show that there are important interactions among the different dimensions, and that these change with age. We also study the dynamics of the process and show that, especially for the case of cognition, a simple Markov structure, with an individual lag of the dimension being studied, might not be enough to capture what is observed in reality.

Consistently with other studies, we have shown that to get consistent estimates of the effect of parental investment on child development, it is important to take into account the endogeneity of such a variable. The impact that parental investments have on child development vary with age and dimension considered. In the case of health, parental investment seems to be important only early on (at age 2). For cognition, instead, parental investment is effective at ages 2,3 and 4. After that, it becomes not particularly important. For socio-emotional skills, instead, parental investment is important for a longer interval of ages and it seems particularly effective at age 5.

We have used the model, which include the interactions among various components, to study how the productivity of investment changes with age, once we take into account these interactions. We also study the sustainability and persistence of certain interventions targeted to different domains of development.

We present a rich analysis of the process of child development which has important implications for policy design. We are not aware of studies that have characterised the process with such high frequency of data over the first period of life. We also introduce some innovations in the way available measures are summarized and how they are compared over different ages.



(a) Increase in investment at age 2

(b) Increase in investment at age 3

(c) Increase in investment at age 4

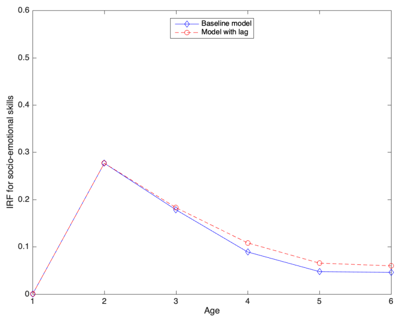
(d) Increase in investment at age 5

(e) Increase in investment at age 2, 3 and 4

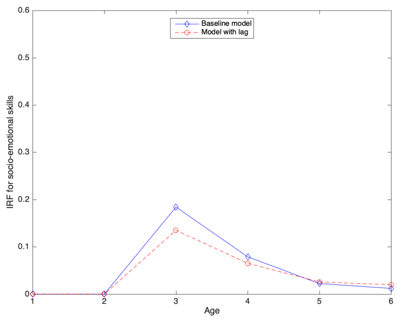
(f) Increase in investment at age 3 and 4

Figure 2.5: IRFs for Cognitive Skills

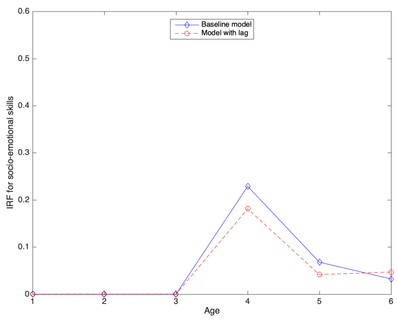
Notes: This figure plots the impulse response functions (IRFs) for cognitive skills following an investment innovation, as implied by the estimated production functions. In each plot, we increase investments by one standard deviation at age 2 (panels *a*), age 3 (panel *b*), age 4 (panel *c*) and age 4 (panel *d*). In panel *e*, investment is increased at ages 2, 3 and 4, while in panel *f*, it is increased at ages 3 and 4. In the plot, we assume that at baseline all other inputs are fixed at their median level. Each graph reports both the evolution of skills predicted by the production function that allows for only one lag of the various dimensions of child development (in blue) and that predicted by the model with an additional self-productivity lag (in red).



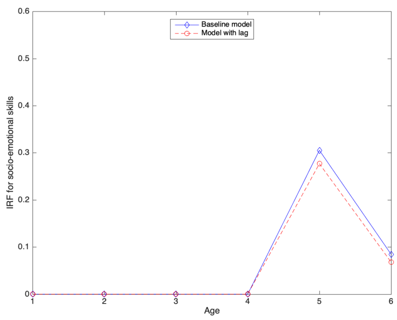
(a) Increase in investment at age 2



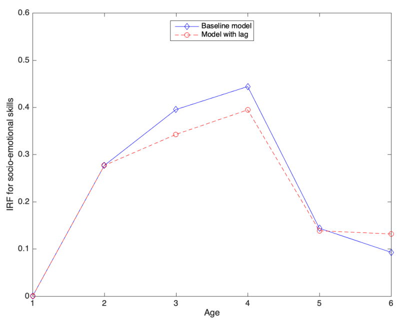
(b) Increase in investment at age 3



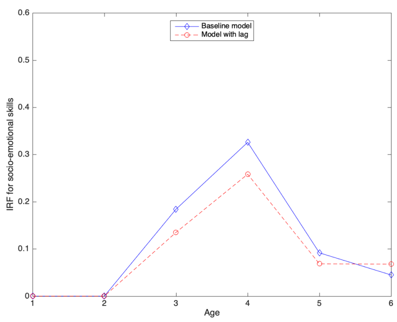
(c) Increase in investment at age 4



(d) Increase in investment at age 5



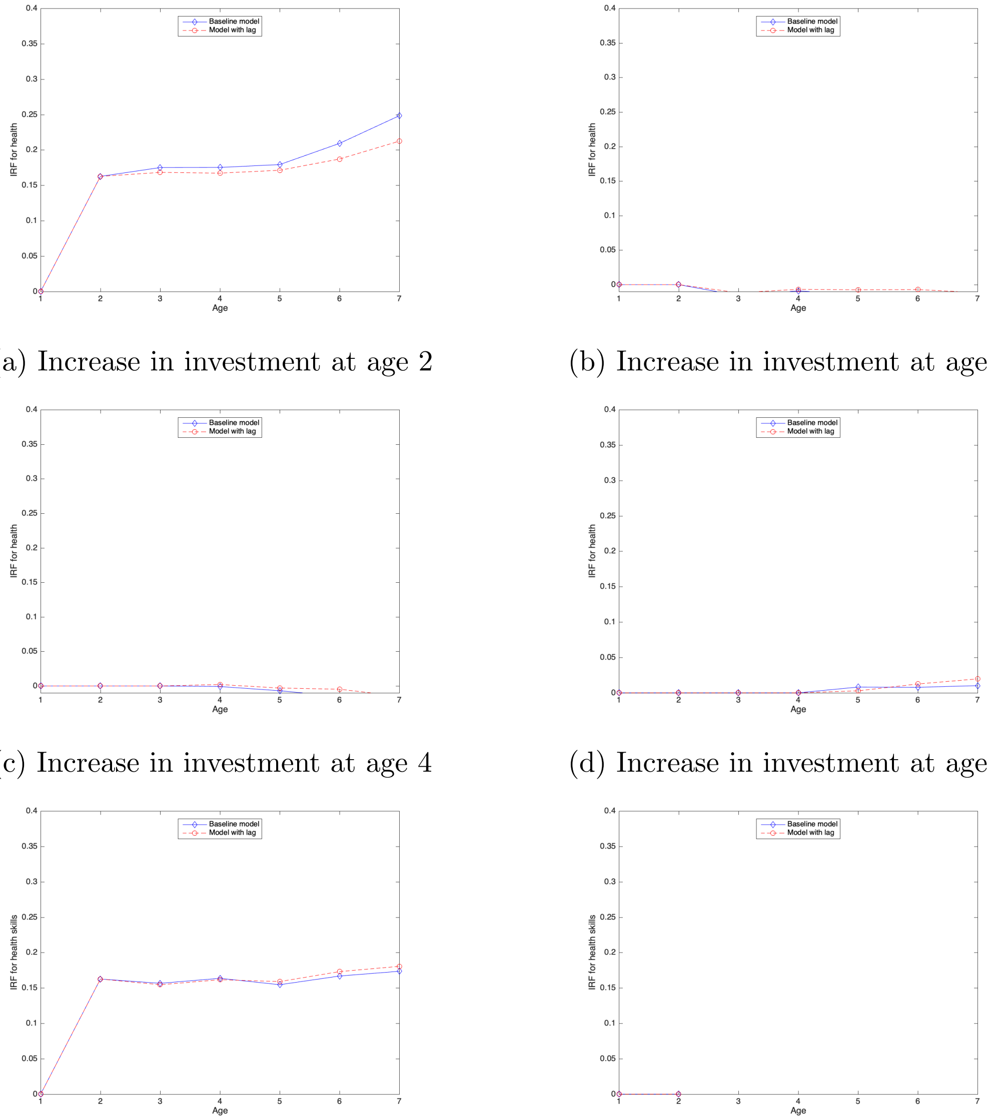
(e) Increase in investment at age 2, 3 and 4



(f) Increase in investment at age 3 and 4

Figure 2.6: IRFs for Socio-emotional Skills

Notes: This figure plots the impulse response functions (IRFs) for socio-emotional skills following an investment innovation, as implied by the estimated production functions. In each plot, we increase investments by one standard deviation at age 2 (panels a), age 3 (panel b), age 4 (panel c) and age 4 (panel d). In panel e, investment is increased at ages 2, 3 and 4, while in panel f, it is increased at ages 3 and 4. In the plot, we assume that at baseline all other inputs are fixed at their median level. Each graph reports both the evolution of skills predicted by the production function that allows for only one lag of the various dimensions of child development (in blue) and that predicted by the model with an additional self-productivity lag (in red).



(a) Increase in investment at age 2

(b) Increase in investment at age 3

(c) Increase in investment at age 4

(d) Increase in investment at age 5

(e) Increase in investment at age 2, 3 and 4

(f) Increase in investment at age 3 and 4

Figure 2.7: IRFs for Health

Notes: This figure plots the impulse response functions (IRFs) for health following an investment innovation, as implied by the estimated production functions. In each plot, we increase investments by one standard deviation at age 2 (panels a), age 3 (panel b), age 4 (panel c) and age 4 (panel d). In panel e, investment is increased at ages 2, 3 and 4, while in panel f, it is increased at ages 3 and 4. In the plot, we assume that at baseline all other inputs are fixed at their median level. Each graph reports both the evolution of skills predicted by the production function that allows for only one lag of the various dimensions of child development (in blue) and that predicted by the model with an additional self-productivity lag (in red).

Appendix 2.A Appendix Tables

Table 2.11: Production Functions for Cognitive Skills (full results)

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.155	-0.135	0.44	0.552	0.74	-0.29	0.027	-0.176	0.481	0.548	0.739
	0.16	0.11	0.03	0.054	0.043	0.05	0.154	0.114	0.044	0.07	0.05	0.05
	[-0.443,0.079]	[0.048,0.402]	[-0.191,-0.098]	[0.341,0.517]	[0.486,0.624]	[0.651,0.811]	[-0.52,-0.037]	[-0.097,0.299]	[-0.278,-0.145]	[0.37,0.588]	[0.456,0.626]	[0.649,0.813]
Control function	-	-	-	-	-	-	-0.22	-0.268	-0.132	0.133	-0.028	-0.061
							0.145	0.102	0.098	0.113	0.161	0.12
							[-0.423,0.036]	[-0.43,-0.127]	[-0.366,-0.053]	[-0.017,0.333]	[-0.272,0.193]	[-0.223,0.158]
AeioTu	-0.13	0.034	0.03	0.068	0.025	0.05	-0.11	0.046	0.03	0.068	0.026	0.056
	0.081	0.056	0.013	0.018	0.045	0.042	0.078	0.051	0.013	0.018	0.046	0.045
	[-0.217,0.038]	[-0.048,0.125]	[0.007,0.047]	[0.026,0.09]	[-0.048,0.099]	[-0.004,0.122]	[-0.208,0.046]	[-0.036,0.129]	[0.007,0.046]	[0.026,0.088]	[-0.051,0.093]	[-0.005,0.134]
Cognitive t	0.3	0.501	0.436	1.021	1.138	0.908	0.3	0.494	0.423	1.051	1.133	0.898
	0.067	0.081	0.039	0.085	0.059	0.048	0.067	0.082	0.042	0.1	0.072	0.054
	[0.229,0.442]	[0.449,0.692]	[0.372,0.498]	[0.879,1.141]	[1.01,1.205]	[0.805,0.954]	[0.221,0.442]	[0.421,0.678]	[0.349,0.487]	[0.901,1.207]	[0.982,1.22]	[0.79,0.969]
Socio-emotional t	0.08	0.158	0.058	-0.06	0.007	0.058	0.08	0.121	0.046	-0.043	0.002	0.053
	0.103	0.075	0.023	0.031	0.036	0.043	0.101	0.074	0.028	0.035	0.049	0.044
	[-0.007,0.335]	[0.018,0.263]	[0.008,0.085]	[-0.089,0.014]	[-0.074,0.044]	[-0.001,0.133]	[-0.014,0.326]	[-0.022,0.215]	[-0.017,0.071]	[-0.071,0.038]	[-0.1,0.063]	[-0.012,0.139]
Health t	0.39	0.238	0.293	0.006	-0.065	-0.062	0.35	0.161	0.267	0.024	-0.066	-0.066
	0.12	0.103	0.048	0.063	0.044	0.034	0.116	0.097	0.055	0.063	0.045	0.035
	[0.134,0.512]	[0.051,0.366]	[0.193,0.362]	[-0.087,0.108]	[-0.119,0.021]	[-0.095,0.011]	[0.108,0.478]	[-0.003,0.319]	[0.156,0.327]	[-0.067,0.137]	[-0.123,0.02]	[-0.103,0.009]
Parental cognition	0.02	0.013	0.102	-0.011	-0.031	0.101	0	-0.008	0.061	0.025	-0.04	0.079
	0.089	0.054	0.031	0.036	0.045	0.045	0.089	0.061	0.044	0.046	0.068	0.064
	[-0.017,0.233]	[-0.011,0.142]	[0.058,0.159]	[-0.051,0.065]	[-0.085,0.064]	[-0.008,0.142]	[-0.047,0.223]	[-0.103,0.063]	[-0.027,0.118]	[-0.023,0.121]	[-0.138,0.086]	[-0.045,0.16]
Investments	0.2	0.091	0.111	0.043	-0.05	-0.005	0.27	0.231	0.203	-0.057	-0.03	0.036
	0.085	0.068	0.041	0.042	0.053	0.045	0.12	0.103	0.092	0.103	0.145	0.097
	[-0.017,0.28]	[-0.035,0.184]	[0.061,0.191]	[-0.06,0.081]	[-0.129,0.045]	[-0.032,0.106]	[-0.032,0.376]	[0.082,0.387]	[0.134,0.418]	[-0.273,0.04]	[-0.234,0.232]	[-0.113,0.19]
Number of children	0.01	0.001	0.005	-0.018	-0.006	-0.001	0.01	0.006	0.008	-0.023	-0.004	0.002
	0.016	0.009	0.006	0.009	0.013	0.011	0.016	0.009	0.007	0.009	0.013	0.011
	[-0.03,0.02]	[-0.014,0.013]	[-0.006,0.012]	[-0.029,0]	[-0.027,0.014]	[-0.017,0.02]	[-0.028,0.02]	[-0.008,0.02]	[-0.003,0.017]	[-0.032,-0.004]	[-0.027,0.015]	[-0.016,0.019]
Gender	0.08	-0.037	-0.037	-0.107	-0.019	0.011	0.07	-0.042	-0.037	-0.108	-0.02	0.006
	0.066	0.046	0.019	0.029	0.032	0.029	0.065	0.042	0.019	0.029	0.034	0.031
	[-0.07,0.146]	[-0.118,0.025]	[-0.06,-0.002]	[-0.137,-0.038]	[-0.071,0.035]	[-0.044,0.052]	[-0.071,0.149]	[-0.114,0.021]	[-0.06,0]	[-0.138,-0.039]	[-0.075,0.034]	[-0.047,0.053]
Age	0	-0.011	0.03	0.059	0.011	-0.037	0	-0.006	0.03	0.059	0.011	-0.035
	0.012	0.011	0.01	0.014	0.015	0.015	0.012	0.011	0.009	0.014	0.015	0.014
	[-0.016,0.023]	[-0.033,0.002]	[0.014,0.045]	[0.033,0.078]	[-0.014,0.034]	[-0.053,-0.009]	[-0.009,0.03]	[-0.025,0.008]	[0.015,0.045]	[0.031,0.076]	[-0.015,0.034]	[-0.052,-0.009]

Notes: This Table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Table 2.12: Production Functions for Cognitive Skills with Two Self-productivity Lags (full results)

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.202	-0.056	0.44	0.422	0.483	-0.29	0.097	-0.098	0.484	0.389	0.478
	0.16	0.102	0.043	0.047	0.073	0.109	0.154	0.101	0.047	0.062	0.083	0.11
	[-0.443,0.079]	[0.1,0.426]	[-0.121,0.012]	[0.353,0.505]	[0.342,0.575]	[0.527,0.876]	[-0.52,-0.037]	[-0.016,0.326]	[-0.199,-0.044]	[0.386,0.574]	[0.298,0.565]	[0.516,0.871]
Control function	-	-	-	-	-	-	-0.22	-0.303	-0.147	0.136	-0.135	-0.077
							0.145	0.107	0.093	0.112	0.157	0.122
							[-0.423,0.036]	[-0.459,-0.124]	[-0.361,-0.066]	[-0.011,0.327]	[-0.395,0.12]	[-0.235,0.145]
AeioTu	-0.13	0.034	0.056	0.068	0.006	0.046	-0.11	0.048	0.057	0.068	0.01	0.054
	0.081	0.057	0.016	0.018	0.046	0.043	0.078	0.051	0.016	0.018	0.047	0.045
	[-0.217,0.038]	[-0.049,0.127]	[0.028,0.08]	[0.025,0.089]	[-0.061,0.085]	[-0.009,0.125]	[-0.208,0.046]	[-0.044,0.127]	[0.028,0.081]	[0.026,0.087]	[-0.055,0.087]	[-0.013,0.132]
Cognitive t	0.3	0.469	0.319	1.022	1.281	1.321	0.3	0.435	0.3	1.044	1.271	1.315
	0.067	0.124	0.069	0.124	0.092	0.178	0.067	0.125	0.071	0.133	0.098	0.181
	[0.229,0.442]	[0.379,0.741]	[0.204,0.427]	[0.808,1.215]	[1.072,1.365]	[0.69,1.211]	[0.221,0.442]	[0.348,0.707]	[0.161,0.403]	[0.834,1.27]	[1.046,1.362]	[0.683,1.213]
Cognitive t-1	-	0.046	0.155	0	-0.248	-0.507	-	0.084	0.162	0.011	-0.275	-0.514
		0.076	0.066	0.073	0.115	0.199		0.075	0.065	0.073	0.117	0.202
		[-0.094,0.134]	[0.067,0.276]	[-0.12,0.101]	[-0.365,0.019]	[-0.377,0.205]		[-0.077,0.147]	[0.071,0.279]	[-0.111,0.121]	[-0.402,-0.014]	[-0.39,0.202]
Socio-emotional t	0.08	0.161	0.049	-0.06	0.053	0.082	0.08	0.123	0.035	-0.043	0.034	0.076
	0.103	0.077	0.023	0.031	0.044	0.042	0.101	0.075	0.027	0.035	0.052	0.043
	[-0.007,0.335]	[0.02,0.264]	[-0.002,0.074]	[-0.088,0.014]	[-0.061,0.085]	[0.006,0.143]	[-0.014,0.326]	[-0.017,0.216]	[-0.027,0.06]	[-0.072,0.039]	[-0.096,0.067]	[-0.012,0.14]
Health t	0.39	0.213	0.269	0.006	-0.03	-0.037	0.35	0.107	0.238	0.02	-0.032	-0.041
	0.12	0.1	0.047	0.051	0.045	0.035	0.116	0.095	0.054	0.053	0.046	0.036
	[0.134,0.512]	[0.034,0.355]	[0.178,0.339]	[-0.066,0.111]	[-0.099,0.044]	[-0.095,0.015]	[0.108,0.478]	[-0.017,0.271]	[0.132,0.309]	[-0.055,0.124]	[-0.105,0.044]	[-0.1,0.018]
Parental cognition	0.02	0.016	0.092	-0.011	0.006	0.12	0	-0.006	0.045	0.026	-0.035	0.092
	0.089	0.059	0.029	0.036	0.048	0.048	0.089	0.062	0.041	0.046	0.067	0.067
	[-0.017,0.233]	[-0.01,0.161]	[0.046,0.144]	[-0.054,0.064]	[-0.065,0.096]	[-0.009,0.149]	[-0.047,0.223]	[-0.092,0.087]	[-0.038,0.094]	[-0.023,0.117]	[-0.14,0.085]	[-0.048,0.171]
Investments	0.2	0.095	0.117	0.043	-0.06	0.02	0.27	0.258	0.221	-0.058	0.036	0.072
	0.085	0.068	0.038	0.044	0.053	0.046	0.12	0.106	0.086	0.103	0.143	0.099
	[-0.017,0.28]	[-0.038,0.176]	[0.066,0.194]	[-0.068,0.083]	[-0.131,0.049]	[-0.036,0.118]	[-0.032,0.376]	[0.088,0.401]	[0.147,0.416]	[-0.275,0.042]	[-0.165,0.297]	[-0.096,0.221]
Number of children	0.01	0	0	-0.018	-0.005	0.001	0.01	0.005	0.004	-0.023	0.003	0.005
	0.016	0.009	0.006	0.009	0.013	0.012	0.016	0.009	0.006	0.009	0.013	0.011
	[-0.03,0.02]	[-0.016,0.013]	[-0.011,0.007]	[-0.029,0]	[-0.027,0.016]	[-0.016,0.023]	[-0.028,0.02]	[-0.008,0.02]	[-0.007,0.012]	[-0.032,-0.004]	[-0.021,0.021]	[-0.014,0.02]
Gender	0.08	-0.039	-0.066	-0.107	0.007	0.016	0.07	-0.045	-0.068	-0.107	0.002	0.009
	0.066	0.046	0.022	0.029	0.032	0.029	0.065	0.042	0.022	0.029	0.034	0.03
	[-0.07,0.146]	[-0.12,0.035]	[-0.093,-0.024]	[-0.141,-0.041]	[-0.059,0.05]	[-0.044,0.051]	[-0.071,0.149]	[-0.112,0.026]	[-0.096,-0.026]	[-0.138,-0.04]	[-0.066,0.05]	[-0.046,0.051]
Age	0	-0.013	0.03	0.059	-0.004	-0.039	0	-0.007	0.029	0.059	-0.004	-0.037
	0.012	0.011	0.01	0.014	0.015	0.015	0.012	0.011	0.01	0.014	0.016	0.014
	[-0.016,0.023]	[-0.033,0.001]	[0.013,0.044]	[0.033,0.077]	[-0.023,0.025]	[-0.053,-0.008]	[-0.009,0.03]	[-0.028,0.005]	[0.014,0.044]	[0.031,0.076]	[-0.023,0.028]	[-0.052,-0.008]

Notes: This Table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Table 2.13: Production Functions for Socio-Emotional Skills (full results)

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.104	0.217	0.312	0.286	0.11	-0.12	0.054	0.09	0.236
	0.303	0.408	0.07	0.064	0.079	0.326	0.409	0.104	0.108	0.089
	[-0.143,0.812]	[-0.596,0.706]	[0.1,0.327]	[0.222,0.451]	[0.158,0.407]	[-0.483,0.597]	[-0.893,0.355]	[-0.141,0.175]	[-0.102,0.251]	[0.102,0.377]
Control function	-	-	-	-	-	-0.62	-0.471	-0.529	-0.728	-0.356
						0.279	0.263	0.237	0.273	0.348
						[-0.979,-0.097]	[-1.045,-0.244]	[-0.955,-0.241]	[-1.288,-0.368]	[-0.947,0.169]
AeioTu	-0.12	0.13	-0.007	-0.115	-0.001	-0.08	0.151	-0.006	-0.117	0.017
	0.181	0.162	0.038	0.031	0.1	0.173	0.156	0.037	0.031	0.106
	[-0.375,0.234]	[-0.305,0.219]	[-0.072,0.05]	[-0.175,-0.072]	[-0.195,0.116]	[-0.343,0.244]	[-0.25,0.218]	[-0.07,0.046]	[-0.172,-0.071]	[-0.201,0.154]
Cognitive t	-0.15	-0.007	0.216	0.099	0.217	-0.16	-0.019	0.167	-0.064	0.151
	0.131	0.206	0.076	0.078	0.088	0.133	0.206	0.088	0.129	0.118
	[-0.369,0.05]	[-0.552,0.138]	[0.062,0.319]	[-0.019,0.224]	[0.072,0.359]	[-0.391,0.035]	[-0.577,0.121]	[-0.019,0.282]	[-0.321,0.083]	[-0.044,0.33]
Socio-emotional t	0.28	0.77	0.41	0.436	0.378	0.26	0.707	0.361	0.345	0.316
	0.159	0.216	0.077	0.067	0.092	0.148	0.229	0.089	0.075	0.109
	[0.118,0.609]	[0.304,0.983]	[0.332,0.568]	[0.332,0.543]	[0.242,0.548]	[0.106,0.538]	[0.208,0.912]	[0.258,0.538]	[0.211,0.454]	[0.148,0.495]
Health t	0.42	0.09	0.156	0.242	0.103	0.31	-0.044	0.05	0.144	0.09
	0.176	0.198	0.082	0.068	0.073	0.184	0.21	0.087	0.086	0.077
	[0.122,0.685]	[-0.106,0.556]	[0.051,0.315]	[0.16,0.38]	[-0.017,0.219]	[0.012,0.605]	[-0.287,0.425]	[-0.077,0.208]	[0.022,0.296]	[-0.069,0.201]
Parental cognition	0.04	-0.021	0.083	0.115	0.206	-0.01	-0.058	-0.084	-0.079	0.088
	0.111	0.129	0.068	0.056	0.089	0.152	0.146	0.095	0.094	0.149
	[-0.071,0.291]	[-0.054,0.37]	[-0.068,0.15]	[0.03,0.217]	[0.054,0.336]	[-0.323,0.145]	[-0.309,0.103]	[-0.309,-0.007]	[-0.231,0.06]	[-0.172,0.287]
Investments	0.41	0.167	0.135	0.108	0.095	0.6	0.415	0.506	0.654	0.354
	0.163	0.137	0.068	0.083	0.099	0.232	0.246	0.203	0.243	0.286
	[0.104,0.593]	[-0.052,0.386]	[0.043,0.264]	[-0.054,0.202]	[-0.049,0.257]	[0.259,0.944]	[0.223,0.989]	[0.256,0.92]	[0.368,1.136]	[-0.049,0.869]
Number of children	0.02	0.003	-0.025	0.026	0	0.02	0.011	-0.01	0.049	0.02
	0.029	0.025	0.018	0.018	0.025	0.027	0.025	0.019	0.019	0.025
	[-0.035,0.057]	[-0.029,0.049]	[-0.053,0.005]	[-0.005,0.051]	[-0.03,0.053]	[-0.021,0.064]	[-0.015,0.066]	[-0.043,0.022]	[0.015,0.075]	[-0.006,0.073]
Gender	0.05	-0.107	-0.072	0.144	-0.15	0.03	-0.114	-0.072	0.146	-0.168
	0.154	0.136	0.06	0.053	0.081	0.147	0.131	0.06	0.054	0.085
	[-0.246,0.254]	[-0.189,0.254]	[-0.17,0.024]	[0.067,0.23]	[-0.257,0.016]	[-0.235,0.254]	[-0.18,0.246]	[-0.173,0.025]	[0.063,0.234]	[-0.278,-0.007]
Age	-0.02	-0.021	0.063	-0.051	0.076	-0.01	-0.011	0.06	-0.049	0.08
	0.029	0.029	0.026	0.027	0.035	0.029	0.029	0.027	0.028	0.035
	[-0.065,0.037]	[-0.093,0]	[0.016,0.098]	[-0.102,-0.014]	[0.01,0.129]	[-0.047,0.045]	[-0.073,0.018]	[0.013,0.1]	[-0.098,-0.01]	[0.011,0.13]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Table 2.14: Production Functions for Socio-emotional Skills with Two Self-productivity Lags (full results)

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.03	0.175	0.316	0.31	0.11	-0.079	0.06	0.107	0.272
	0.303	0.535	0.076	0.065	0.078	0.326	0.508	0.102	0.106	0.088
	[-0.143,0.812]	[-0.997,0.783]	[0.034,0.271]	[0.226,0.438]	[0.195,0.452]	[-0.483,0.597]	[-1.126,0.534]	[-0.133,0.177]	[-0.053,0.268]	[0.159,0.424]
Control function	-	-	-	-	-	-0.62	-0.258	-0.447	-0.683	-0.242
						0.279	0.243	0.25	0.268	0.347
						[-0.979,-0.097]	[-0.747,0.012]	[-0.935,-0.127]	[-1.228,-0.303]	[-0.749,0.369]
AeioTu	-0.12	0.147	0.035	-0.1	0.01	-0.08	0.158	0.012	-0.105	0.021
	0.181	0.15	0.038	0.033	0.097	0.173	0.148	0.039	0.034	0.102
	[-0.375,0.234]	[-0.222,0.24]	[-0.034,0.093]	[-0.158,-0.05]	[-0.18,0.125]	[-0.343,0.244]	[-0.217,0.271]	[-0.058,0.062]	[-0.16,-0.059]	[-0.191,0.142]
Cognitive t	-0.15	-0.041	0.212	0.093	0.17	-0.16	-0.04	0.173	-0.059	0.133
	0.131	0.272	0.077	0.081	0.094	0.133	0.266	0.088	0.124	0.116
	[-0.369,0.05]	[-0.721,0.182]	[0.051,0.31]	[-0.032,0.23]	[-0.003,0.312]	[-0.391,0.035]	[-0.716,0.138]	[-0.018,0.282]	[-0.286,0.1]	[-0.064,0.313]
Socio-emotional t	0.28	0.728	0.39	0.338	0.314	0.26	0.702	0.36	0.275	0.282
	0.159	0.211	0.085	0.076	0.103	0.148	0.224	0.09	0.078	0.111
	[0.118,0.609]	[0.245,0.918]	[0.293,0.557]	[0.193,0.448]	[0.149,0.484]	[0.106,0.538]	[0.173,0.887]	[0.248,0.534]	[0.116,0.369]	[0.099,0.46]
Socio-emotional t-1	-	0.087	0.171	0.114	0.164	-	0.066	0.075	0.088	0.139
		0.246	0.093	0.079	0.119		0.246	0.096	0.078	0.125
		[-0.265,0.61]	[0.034,0.3]	[0.054,0.283]	[0.013,0.393]		[-0.243,0.588]	[-0.081,0.215]	[0.015,0.242]	[-0.024,0.383]
Health t	0.42	0.078	0.118	0.263	0.093	0.31	0.005	0.05	0.166	0.086
	0.176	0.203	0.083	0.071	0.072	0.184	0.21	0.085	0.087	0.074
	[0.122,0.685]	[-0.121,0.555]	[0.003,0.268]	[0.162,0.397]	[-0.025,0.205]	[0.012,0.605]	[-0.246,0.468]	[-0.073,0.203]	[0.031,0.318]	[-0.043,0.196]
Parental cognition	0.04	-0.024	0.065	0.111	0.178	-0.01	-0.044	-0.066	-0.07	0.102
	0.111	0.112	0.066	0.055	0.088	0.152	0.122	0.097	0.09	0.144
	[-0.071,0.291]	[-0.097,0.249]	[-0.084,0.13]	[0.016,0.203]	[0.007,0.302]	[-0.323,0.145]	[-0.256,0.101]	[-0.28,0.018]	[-0.227,0.062]	[-0.148,0.312]
Investments	0.41	0.173	0.044	0.082	0.081	0.6	0.311	0.409	0.6	0.259
	0.163	0.15	0.071	0.089	0.095	0.232	0.22	0.22	0.247	0.28
	[0.104,0.593]	[-0.029,0.45]	[-0.042,0.199]	[-0.101,0.171]	[-0.067,0.245]	[0.259,0.944]	[0.087,0.843]	[0.204,0.895]	[0.222,1.072]	[-0.189,0.697]
Number of children	0.02	-0.004	-0.032	0.02	0.007	0.02	0.002	-0.016	0.042	0.02
	0.029	0.031	0.018	0.018	0.026	0.027	0.031	0.018	0.018	0.023
	[-0.035,0.057]	[-0.061,0.039]	[-0.059,0.001]	[-0.011,0.045]	[-0.021,0.064]	[-0.021,0.064]	[-0.045,0.049]	[-0.043,0.015]	[0.005,0.069]	[-0.004,0.064]
Gender	0.05	-0.122	-0.105	0.117	-0.143	0.03	-0.125	-0.087	0.125	-0.157
	0.154	0.123	0.059	0.055	0.077	0.147	0.122	0.058	0.056	0.08
	[-0.246,0.254]	[-0.203,0.177]	[-0.201,-0.009]	[0.037,0.21]	[-0.243,-0.003]	[-0.235,0.254]	[-0.189,0.178]	[-0.182,0.015]	[0.042,0.215]	[-0.253,-0.007]
Age	-0.02	-0.02	0.075	-0.03	0.075	-0.01	-0.014	0.066	-0.033	0.077
	0.029	0.031	0.024	0.029	0.034	0.029	0.03	0.026	0.03	0.033
	[-0.065,0.037]	[-0.092,0.013]	[0.027,0.107]	[-0.08,0.013]	[0.01,0.124]	[-0.047,0.045]	[-0.079,0.017]	[0.023,0.102]	[-0.08,0.015]	[0.015,0.124]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Table 2.15: Production Functions for Health (full results)

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.382	0.347	0.38	0.331	0.496	-0.06	0.402	0.368	0.353	0.35	0.497
	0.145	0.062	0.017	0.021	0.026	0.026	0.135	0.071	0.021	0.025	0.033	0.027
	[-0.14,0.319]	[0.252,0.449]	[0.323,0.379]	[0.338,0.407]	[0.305,0.391]	[0.455,0.538]	[-0.288,0.166]	[0.235,0.47]	[0.341,0.409]	[0.307,0.391]	[0.323,0.425]	[0.455,0.539]
Control function	-	-	-	-	-	-	-0.33	0.042	0.067	-0.09	0.132	0.033
							0.116	0.061	0.037	0.046	0.137	0.098
							[-0.473,-0.119]	[-0.102,0.106]	[0.015,0.128]	[-0.13,0.011]	[-0.012,0.433]	[-0.061,0.236]
AeioTu	-0.03	0.004	0.005	0.005	0.069	0.105	0	0.002	0.005	0.005	0.062	0.102
	0.061	0.042	0.005	0.007	0.032	0.028	0.05	0.041	0.005	0.007	0.032	0.03
	[-0.104,0.086]	[-0.069,0.059]	[-0.002,0.016]	[-0.009,0.014]	[0.004,0.105]	[0.05,0.137]	[-0.076,0.076]	[-0.068,0.06]	[-0.002,0.016]	[-0.01,0.014]	[-0.005,0.097]	[0.04,0.133]
Cognitive t	-0.02	0.061	0.002	-0.13	-0.1	-0.122	-0.03	0.062	0.008	-0.151	-0.075	-0.117
	0.063	0.1	0.022	0.033	0.038	0.033	0.062	0.098	0.022	0.036	0.049	0.036
	[-0.108,0.096]	[-0.043,0.268]	[-0.037,0.037]	[-0.191,-0.086]	[-0.152,-0.031]	[-0.177,-0.072]	[-0.108,0.093]	[-0.043,0.26]	[-0.03,0.044]	[-0.214,-0.093]	[-0.126,0.026]	[-0.171,-0.05]
Socio-emotional t	-0.01	0.069	0.026	0.042	-0.037	-0.035	-0.02	0.075	0.032	0.031	-0.014	-0.032
	0.098	0.052	0.011	0.014	0.036	0.027	0.087	0.05	0.014	0.016	0.049	0.028
	[-0.1,0.225]	[0.003,0.18]	[0.006,0.041]	[0.009,0.055]	[-0.088,0.028]	[-0.085,0.005]	[-0.1,0.187]	[0.006,0.169]	[0.012,0.057]	[-0.005,0.052]	[-0.07,0.079]	[-0.078,0.021]
Health t	0.77	0.892	0.95	1.107	1.196	1.237	0.71	0.904	0.963	1.095	1.202	1.239
	0.107	0.155	0.028	0.034	0.043	0.034	0.097	0.162	0.03	0.034	0.045	0.036
	[0.533,0.872]	[0.543,1.036]	[0.914,1.009]	[1.051,1.164]	[1.103,1.249]	[1.173,1.291]	[0.495,0.822]	[0.522,1.056]	[0.925,1.022]	[1.047,1.159]	[1.11,1.263]	[1.181,1.301]
Parental cognition	0	-0.006	-0.023	0.029	-0.017	-0.071	-0.03	-0.002	-0.002	0.005	0.027	-0.06
	0.071	0.029	0.015	0.02	0.035	0.032	0.058	0.031	0.018	0.023	0.063	0.049
	[-0.017,0.189]	[-0.026,0.062]	[-0.047,0.003]	[-0.013,0.052]	[-0.075,0.031]	[-0.123,-0.017]	[-0.062,0.095]	[-0.028,0.062]	[-0.03,0.03]	[-0.04,0.036]	[-0.049,0.16]	[-0.118,0.034]
Investments	0.27	-0.017	0.045	-0.048	-0.043	-0.009	0.37	-0.039	-0.002	0.02	-0.139	-0.031
	0.073	0.045	0.019	0.022	0.04	0.035	0.083	0.064	0.034	0.041	0.13	0.082
	[0.061,0.301]	[-0.061,0.082]	[0.01,0.069]	[-0.059,0.014]	[-0.102,0.023]	[-0.053,0.061]	[0.18,0.443]	[-0.117,0.099]	[-0.063,0.046]	[-0.038,0.101]	[-0.401,0.006]	[-0.196,0.061]
Number of children	0	0.001	0.002	-0.011	-0.01	-0.013	0.01	0	0	-0.008	-0.018	-0.014
	0.013	0.005	0.003	0.004	0.007	0.008	0.013	0.006	0.003	0.003	0.009	0.007
	[-0.022,0.019]	[-0.007,0.011]	[-0.002,0.007]	[-0.013,-0.001]	[-0.021,0]	[-0.024,0.001]	[-0.017,0.023]	[-0.008,0.012]	[-0.004,0.006]	[-0.011,0]	[-0.034,-0.006]	[-0.027,-0.004]
Gender	0.01	0.012	0.011	-0.011	-0.058	-0.085	0	0.013	0.011	-0.011	-0.051	-0.082
	0.049	0.037	0.009	0.011	0.026	0.019	0.042	0.037	0.009	0.011	0.028	0.02
	[-0.081,0.073]	[-0.039,0.076]	[-0.008,0.023]	[-0.023,0.013]	[-0.086,0.001]	[-0.105,-0.043]	[-0.07,0.061]	[-0.04,0.076]	[-0.009,0.024]	[-0.023,0.013]	[-0.076,0.01]	[-0.1,-0.034]
Age	0.01	0	-0.004	0.007	0.028	-0.016	0.01	-0.001	-0.003	0.007	0.026	-0.017
	0.009	0.006	0.005	0.006	0.011	0.009	0.01	0.006	0.005	0.006	0.012	0.009
	[-0.009,0.02]	[-0.006,0.016]	[-0.012,0.004]	[-0.005,0.012]	[0.002,0.042]	[-0.032,-0.004]	[-0.001,0.029]	[-0.006,0.015]	[-0.011,0.004]	[-0.005,0.012]	[-0.002,0.039]	[-0.035,-0.006]

*Notes:*This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Table 2.16: Production Functions for Health with Two Self-Productivity Lags (full results)

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.394	0.298	0.256	0.076	0.376	-0.06	0.407	0.318	0.222	0.08	0.371
	0.145	0.065	0.026	0.072	0.095	0.059	0.135	0.068	0.028	0.079	0.097	0.062
	[-0.14,0.319]	[0.252,0.472]	[0.261,0.347]	[0.298,0.53]	[-0.022,0.28]	[0.306,0.486]	[-0.288,0.166]	[0.249,0.479]	[0.282,0.369]	[0.266,0.517]	[-0.003,0.3]	[0.301,0.498]
Control function	-	-	-	-	-	-	-0.33	0.034	0.06	-0.1	0.204	-0.032
							0.116	0.07	0.038	0.048	0.12	0.099
							[-0.473,-0.119]	[-0.112,0.101]	[0.002,0.12]	[-0.13,0.018]	[0.027,0.411]	[-0.125,0.184]
AeioTu	-0.03	0.004	0.009	0.001	0.028	0.074	0	0.003	0.009	0.001	0.014	0.076
	0.061	0.042	0.006	0.007	0.03	0.029	0.05	0.042	0.006	0.007	0.03	0.03
	[-0.104,0.086]	[-0.068,0.058]	[0.002,0.02]	[-0.008,0.014]	[-0.026,0.068]	[0.029,0.121]	[-0.076,0.076]	[-0.069,0.059]	[0.002,0.021]	[-0.008,0.014]	[-0.034,0.061]	[0.027,0.119]
Cognitive t	-0.02	0.05	0.012	-0.092	-0.032	-0.092	-0.03	0.054	0.017	-0.113	0.012	-0.096
	0.063	0.084	0.023	0.042	0.041	0.033	0.062	0.082	0.023	0.043	0.048	0.036
	[-0.108,0.096]	[-0.069,0.212]	[-0.031,0.047]	[-0.22,-0.077]	[-0.091,0.042]	[-0.157,-0.045]	[-0.108,0.093]	[-0.062,0.203]	[-0.025,0.054]	[-0.24,-0.094]	[-0.054,0.095]	[-0.158,-0.044]
Socio-emotional t	-0.01	0.067	0.025	0.038	-0.007	-0.01	-0.02	0.072	0.031	0.025	0.031	-0.012
	0.098	0.051	0.012	0.014	0.039	0.029	0.087	0.049	0.014	0.017	0.045	0.028
	[-0.1,0.225]	[0.002,0.168]	[0.005,0.041]	[0.008,0.056]	[-0.058,0.063]	[-0.068,0.032]	[-0.1,0.187]	[0.005,0.153]	[0.008,0.053]	[-0.004,0.054]	[-0.03,0.113]	[-0.063,0.028]
Health t	0.77	0.857	1.051	1.472	1.752	1.546	0.71	0.876	1.06	1.47	1.815	1.554
	0.107	0.233	0.066	0.215	0.208	0.152	0.097	0.248	0.067	0.217	0.206	0.16
	[0.533,0.872]	[0.381,1.119]	[0.955,1.175]	[0.639,1.327]	[1.278,1.97]	[1.26,1.744]	[0.495,0.822]	[0.334,1.132]	[0.965,1.186]	[0.654,1.335]	[1.323,2.006]	[1.226,1.743]
Health t-1	-	0.041	-0.117	-0.385	-0.698	-0.396	-	0.03	-0.113	-0.397	-0.766	-0.409
		0.144	0.06	0.223	0.252	0.176		0.149	0.061	0.228	0.249	0.191
		[-0.178,0.272]	[-0.214,-0.021]	[-0.219,0.491]	[-0.976,-0.155]	[-0.626,-0.066]		[-0.192,0.286]	[-0.208,-0.02]	[-0.229,0.488]	[-1.005,-0.193]	[-0.635,-0.018]
Parental cognition	0	-0.006	-0.019	0.034	0.012	-0.063	-0.03	-0.003	0	0.007	0.082	-0.074
	0.071	0.033	0.016	0.02	0.036	0.033	0.058	0.032	0.019	0.023	0.056	0.045
	[-0.017,0.189]	[-0.026,0.079]	[-0.041,0.011]	[-0.015,0.05]	[-0.053,0.056]	[-0.116,-0.008]	[-0.062,0.095]	[-0.028,0.063]	[-0.027,0.034]	[-0.038,0.036]	[-0.015,0.172]	[-0.119,0.014]
Investments	0.27	-0.009	0.047	-0.068	-0.027	0.015	0.37	-0.029	0.005	0.007	-0.174	0.036
	0.073	0.053	0.019	0.028	0.037	0.038	0.083	0.081	0.034	0.042	0.109	0.081
	[0.061,0.301]	[-0.069,0.095]	[0.01,0.073]	[-0.065,0.025]	[-0.089,0.028]	[-0.039,0.076]	[0.18,0.443]	[-0.119,0.141]	[-0.059,0.053]	[-0.035,0.107]	[-0.382,-0.037]	[-0.122,0.122]
Number of children	0	0.001	0.002	-0.012	0	-0.009	0.01	0	0.001	-0.009	-0.011	-0.007
	0.013	0.005	0.003	0.004	0.007	0.008	0.013	0.006	0.003	0.004	0.008	0.007
	[-0.022,0.019]	[-0.008,0.011]	[-0.002,0.007]	[-0.013,0]	[-0.012,0.009]	[-0.019,0.005]	[-0.017,0.023]	[-0.007,0.012]	[-0.004,0.006]	[-0.011,0.001]	[-0.026,0]	[-0.021,0.002]
Gender	0.01	0.011	0.005	-0.012	-0.028	-0.055	0	0.012	0.005	-0.012	-0.015	-0.057
	0.049	0.036	0.009	0.011	0.024	0.02	0.042	0.036	0.009	0.011	0.025	0.02
	[-0.081,0.073]	[-0.042,0.075]	[-0.013,0.019]	[-0.024,0.013]	[-0.059,0.015]	[-0.084,-0.02]	[-0.07,0.061]	[-0.041,0.075]	[-0.014,0.019]	[-0.024,0.012]	[-0.051,0.031]	[-0.084,-0.02]
Age	0.01	0	0	0.009	0.025	-0.028	0.01	-0.001	0	0.01	0.023	-0.028
	0.009	0.006	0.005	0.006	0.011	0.009	0.01	0.006	0.005	0.006	0.012	0.009
	[-0.009,0.02]	[-0.006,0.016]	[-0.008,0.009]	[-0.006,0.012]	[0.003,0.039]	[-0.039,-0.011]	[-0.001,0.029]	[-0.006,0.015]	[-0.008,0.009]	[-0.006,0.012]	[-0.003,0.038]	[-0.041,-0.01]

Notes: This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

Chapter 3

Parental Investments and Intra-household Inequality in Child Development: Theory, Measurement, and Evidence from a Lab-in-the-Field Experiment

3.1 Introduction

Intra-household inequality is key for the measurement and understanding of poverty and inequality (Chiappori & Meghir (2015); Brown, Ravallion, & Van De Walle (2017)). This paper focuses on inequality in human capital outcomes *between* children in the same family. Because the early years are critical for the process of human capital formation and can have important consequences for well-being through the life cycle (Currie & Almond (2011); Currie & Vogl (2013); Brito & Noble (2014); Heckman & Mosso (2014)), understanding the origin of this inequality is key to the design of effective policies that aim at reducing disparities across individuals. In particular, if intra-household inequality between siblings is driven by parental investment choices, understanding what determines these choices can help policy-makers design better interventions that account for parents' endogenous responses.[1]

To study the sources of intra-household inequality in child outcomes, I combine a theoretical model of household behaviour with experimental data from India. I use the theoretical framework to motivate the empirical analyses, design the survey strategy and measurement tool used in the field, and interpret the empirical findings. The model fleshes out the separate role that parental preferences, beliefs and resource constraints have in determining educational investments, highlighting the challenge

[1] Relatedly, recent work has documented that intra-household inequality between children and adults has important consequences for child poverty (Dunbar, Lewbel, & Pendakur (2013)).

that the use of observational data poses to the identification of key parameters and mechanisms of interest. The reason for this challenge is a twofold identification issue. First, realized choices may be consistent with many alternative specifications of preferences and beliefs (Manski (2004)). Second, resource constraints might prevent parents from selecting their preferred choices, breaking the direct connection between observed outcomes and parental preferences. Guided by this theoretical framework, I design and implement a lab-in-the-field experiment that allows me to overcome these identification challenges. I develop a novel survey methodology based on hypothetical scenarios to elicit direct measures of parental beliefs, identify preferences for intra-household inequality, and study the role that resource constraints have to determine parental investments in their children's education. I then show that primitive parameters identified in the experiment are predictive of *actual* parental behaviour outside the experiment.

The theoretical analysis builds on the conventional literature on intra-household resource allocation. In the model parents choose how to allocate educational investments between their children. Because of limited resources, they face a trade off between *efficiency* and *equity* considerations. Inequality arises through two distinct channels. The first is a direct channel operating through the production function for human capital. Children have different early levels of human capital – or endowments – and these endowments have a direct effect on child human capital (Currie & Hyson (1999); Behrman & Rosenzweig (2004b); Royer (2009); Almond & Mazumder (2011); Figlio, Guryan, Karbownik, & Roth (2014)). The second indirect channel operates through parental investments. By differentially allocating resources between their children, in ways that reinforce or compensate initial differences, parents may exacerbate or attenuate the effect of endowments (Becker & Tomes (1976); Behrman, Pollak, & Taubman (1982); Behrman (1988); Pitt, Rosenzweig, & Hassan (1990)).

I extend this analysis in two ways. First, I allow parents to have inaccurate beliefs about the production function for child human capital. The standard practice in economics to investigate what drives parental allocations, is to focus on the interplay between the technology of skill formation and parental preferences for intra-household inequality in child human capital outcomes. The implicit assumption underlying these models is that parents have full knowledge of the process of child development when making these decisions. Previous work has therefore relied on investment data to derive conclusions on the nature of parental preferences (Behrman, Pollak, & Taubman (1982)). Given the great difficulties in identifying the properties of the human capital production function, it seems likely that investment choices are made under imperfect information. Indeed, recent evidence suggests that parents hold inaccurate beliefs about the productivity of different inputs entering the production function for child human capital, and are mistaken about important features of the process of human capital accumulation (Boneva & Rauh (2018); Attanasio, Cunha, & Jervis (2019)).[2]

[2]There is also evidence that parents might have inaccurate perceptions about their children's

Second, I explicitly study how household resources affect allocation choices. This might be particularly important in poor or large families, where per-capita resources are lower.[3] I show that if children have heterogeneous endowments, an increase in family size will result in lower per-capita resources, and may lead parents to differentially invest in their children, increasing the overall level of inequality in the family.

By incorporating these frictions into a standard model of intra-household allocation of resources, I show that identification issues are of first order importance: Beliefs and constraints are usually not observed in standard survey data. Therefore observational data on parental allocative choices do not allow to separate their role from that of preferences. Clearly, these distinctions are crucial to understand the sources of intra-household inequality, and to the design of policies that could effectively tackle it.

Guided by this theoretical framework, I design a novel measurement strategy based on hypothetical scenarios that allows me to elicit parental beliefs about the human capital production function, and identify preferences for intra-household inequality. I embed this measurement tool in a lab-in-the-field experiment with parents of primary school children in the slums of a large city in India. The experiment consists of two stages. I the first stage, I identify parental beliefs about the human capital production function. The approach used to elicit these beliefs builds on the work by Cunha, Elo, & Culhane (2013), as recently extended by Boneva & Rauh (2018) and Attanasio, Boneva, & Rauh (2019). It consists in presenting a series of hypothetical stories to the respondent and elicit information on individual expected outcomes. By varying the characteristics of the scenarios one at a time while keeping other factors constant, one can trace out the perceived human capital production function. Following the theoretical model, I focus on the role of child endowments (captured by academic ability) and parental educational investments.

Having identified parental beliefs about the human capital production function, in the second stage of the experiment I collect parents' stated investment choices. As in the case of beliefs, respondents are presented with 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, I ask them to select their favourite allocation choice. The design of the scenarios closely follows the theoretical framework. In particular, because the model highlights that investments might depend on child ability and household resources, scenarios vary according to these two key dimensions.

Combining these strategical designed survey measures on beliefs and choices, I identify parents' preferences at the time of the survey, free from other confounding

---

human capital, causing the mis-allocation of education investments (Dizon-Ross (2019)). The empirical strategy that I use in this paper is robust to the possibility that parents might have inaccurate beliefs about their children's human capital.

[3]The idea that parental investments might be affected by family size goes back to the seminal Quantity-Quality (Q-Q) model by Becker & Lewis (1973), which predicts that decreases in fertility will induce more resources to be allocated to each child so that average levels of child human capital will increase. Importantly, this model rests on the implicit assumption that the quality of each child is the same.

factors. Importantly, because scenarios vary in terms of resources available to the family, I can also study the role that resource constraints have in determining investments. By directly eliciting information about the *perceived* production function, I can avoid making strong assumptions on parental information sets and beliefs, upon which earlier work relies.[4]

Several key results emerge from this study. First, I find that parents perceive ability and investments to be highly productive. A one-standard-deviation increase in endowments is perceived to increase earnings by 15 percent; a similar increase in investments boosts adult earnings by 28 percent. Moreover, parents perceive investments and endowments to be complements i.e. they believe that investments are more productive for higher ability children. This perceived complementarity generates an incentive for parents to reinforce initial differences across their children. By showing that parental beliefs about the human capital production function matter for intra-household allocations, I contribute to a growing literature focusing on the role of subjective beliefs as a determinant of human capital investment decisions.[5] I advance this literature by documenting that perceived returns matter to explain differences in investments *within* a family across children, beside their importance to explain inequalities *between* families (Boneva & Rauh (2018); Attanasio, Cunha, & Jervis (2019)).

Second, the experimental results reveal that parents are not averse to inequality over their children's human capital outcomes and reinforce differences in child endowments. Specifically, I show that when the difference in children's ability increases, parents re-allocate resources towards the higher achieving child, suggesting that in this setting parental investment choices are to some extent driven by efficiency considerations (Becker & Tomes (1976); Griliches (1979); Behrman, Pollak, & Taubman (1982); Behrman (1988); Pitt, Rosenzweig, & Hassan (1990)).

Third, I show that household resources are important in explaining investment choices. In particular, I find that parents reinforce more strongly initial conditions

[4]In independent work developed contemporaneously to this paper Berry, Dizon-Ross, & Jagnani (2020) (BDJ) also use a lab-in-the-field experiment to study parental preferences for investing in their children. There are a few important differences between BDJ and this paper. First, rather than using data on parental beliefs about the production function, BDJ shocks to the short run returns to investing in different children for identification. Each of these two identification strategies has advantages and disadvantages. On the one hand, the strategy in BDJ does not allow studying the role that individual perceived returns have in explaining household behaviour. Moreover, changing the child specific payments functions that map test scores to final outcomes might affect parents' choices, as the way parents make investment decisions might be affected by the environment where the decisions are made (e.g. an environment that highly rewards test score points might give an incentive for parents to reward the higher achieving child). On the other hand, controlling the function that maps choices to final outcomes, allows BDJ to consider the possibility that parents have a preferences for equalizing inputs, rather than final outcomes. This is not possible in the context of the current experiment. Second, BDJ do not study the role of household resources. As I discuss later, these play a key role to explain parental investments.

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

and select more unequal allocations when resources are lower. This result adds to and complements a large literature investigating the role that credit constraints have to explain socio-economic gaps in school enrolment and educational investments (Lochner & Monge-Naranjo (2012); Kaufmann (2014); Solis (2017)). I contribute to this literature, by presenting evidence that resource constraints have important implications for the *allocation* of human capital investments across siblings.

More broadly, by demonstrating how factors that have been shown to matter for inequality in child outcomes *across* households are also key to explain inequality within the family, this paper relates to a growing body of evidence pointing at the importance of considering intra-household inequality to understand differences across individuals in a society.[6] While the traditional focus of this literature is the measurement of inequality across different groups of individuals living in the household (e.g. men *vs.* women; adults *vs.* children), I document the importance of intra-household inequality between siblings' human capital outcomes and propose a mechanism that explains the existence of this inequality. Both the model and the empirical findings suggest that child endowments might have an important role in driving inequality in investments and child outcomes.

Finally, in terms of field methodology, this paper relates to a growing literature using hypothetical scenarios to collect data on individual beliefs, and elicit stated choices to understand behaviour.[7] I show how strategically designed, theory-driven survey measures can be combined to identify primitive parameters of interest. One implicit assumption about this methodology is that stated choices reported in the experiment are reflective of what respondents would do in actual scenarios, i.e. that elicited preferences relate to *actual* behaviour. 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 respondents (Mas & Pallais (2017); Wiswall & Zafar (2018)). To address this question, I also collect data on *actual* educational investments made by study participants in their own children and find a strong relation between elicited preferences and realized investments. In particular, respondents that are identified as less inequality averse in the experiment, spend more unequally on their children's education. I also find that higher endowment children are more likely to attend a private school and less likely to work, suggesting that these decision extend beyond educational expenditure at a particular point in time to higher stakes investment

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

[7]Cunha, Elo, & Culhane (2013), Attanasio, Boneva, & Rauh (2019), Boneva & Rauh (2018), Attanasio, Cunha, & Jervis (2019) elicit data on parental beliefs about the child human capital production function. Importantly, these papers do not use these data to identify parental preferences. Mas & Pallais (2017), Wiswall & Zafar (2018), Adams & Andrew (2019), Ameriks, Briggs, Caplin, Lee, et al. (2020), Ameriks, Briggs, Caplin, Shapiro, & Tonetti (2020), use stated choice data to study preferences for workplace attributes, university choices, marriage markets, saving behaviour and labour force participation.

choices that can have potentially important long-term effects. These results add credibility to the research design and to the use of hypothetical scenarios to identify key parameters of interest.

The remainder of the paper is organized as follows. In the following section, I present some basic stylized facts that provide the motivation for the study. Section 3.3 present a simple conceptual framework that can be used to study intra-household inequality in child outcomes. Section 3.4 describes the experiment. Section 3.5 describes the setting and the data. I discuss the results in Section 3.6 and conclude in Section 3.7.

3.2 Motivating Evidence

The basic empirical evidence motivating this study is presented in Figure 3.1. This figure plots the shares of total variation in child educational attainment that can be attributed to the within- and between-households components. 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 and a between-groups components.[8] The figure shows that, across developing countries, intra-household variation explains between 30 to 45 percent of overall inequality in child human capital.

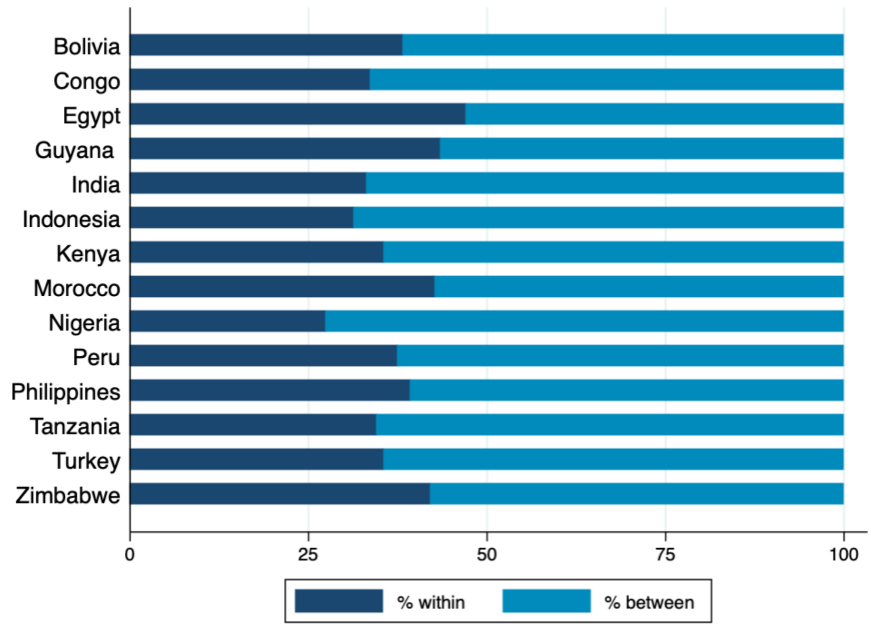


Figure 3.1: Inequality in Child Human Capital

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. Thus coefficients are expressed in standard deviations units. Each bar represents a different country. Source: Development and Health Survey (DHS).

[8]See Appendix 3.A for details about the MLD measure of inequality, and its decomposition.

In India, the country under study in this paper, inequality between siblings amounts to 33 percent of overall inequality in outcomes.[9]

In panel A of Figure 3.2, I plot the same relation but stratifying the sample by family size. In particular, the figure shows the intra-household contribution to overall inequality separately for families with a different number of children. The figure reveals that the share of variance in child human capital explained by within-household variation rises from 25 percent in a two-children family to 70 percent in families with six or more children. To study what explains this large increase, in panel B of Figure 3.2 I plot the human capital distribution by family size, focusing on the mean, the maximum, and the minimum of the distribution (i.e. the human capital of the highest and lowest-achieving child, and the average level of human capital in the family). The figure reveals several interesting patterns.

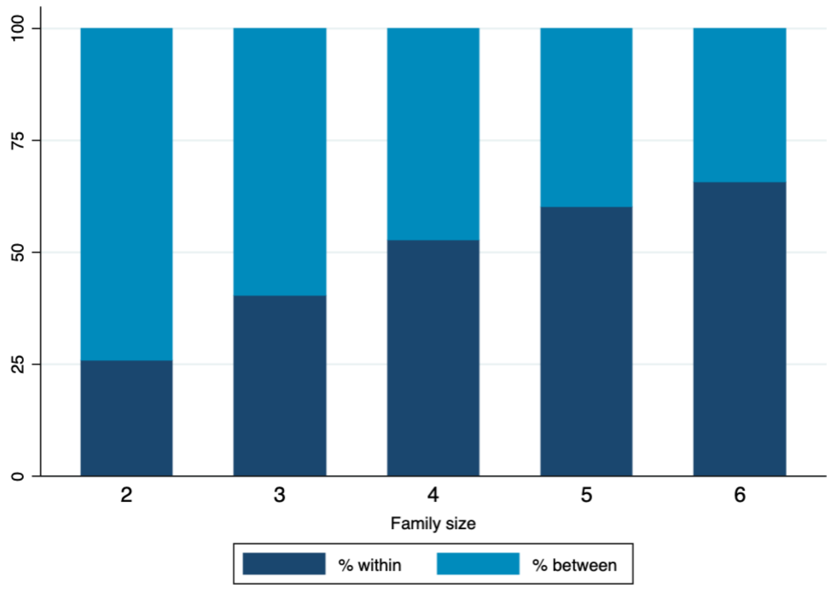
First, there is a negative relationship between average child quality and family size. This is a relatively well know fact that goes back to Becker's Q-Q model, and can be explained by the fact that there are less per-capita resources in larger families, so that each child will receive less educational investments (Becker & Lewis (1973)). More interestingly, the human capital of the most successful child in the family varies vary little with family size. On the other hand, the human capital of the least successful child steeply declines as family size increases.[10,11]

These patterns are interesting as they provide information on the origins of human capital. They indicate that parental investment might play an important role to explain intra-household inequality. Specifically, these patterns are consistent with the idea that resource constrained families might distribute resources unequally across their children and focus their educational investments on one child in the family. This, in turn, has a large detrimental effect on the human capital of other children in the household, particularly in larger families, where per-capita resources are lower. Clearly, this investment strategy crucially depends on parental preferences for intra-household inequality in child outcomes. A key contribution of this paper is to identify these preferences.

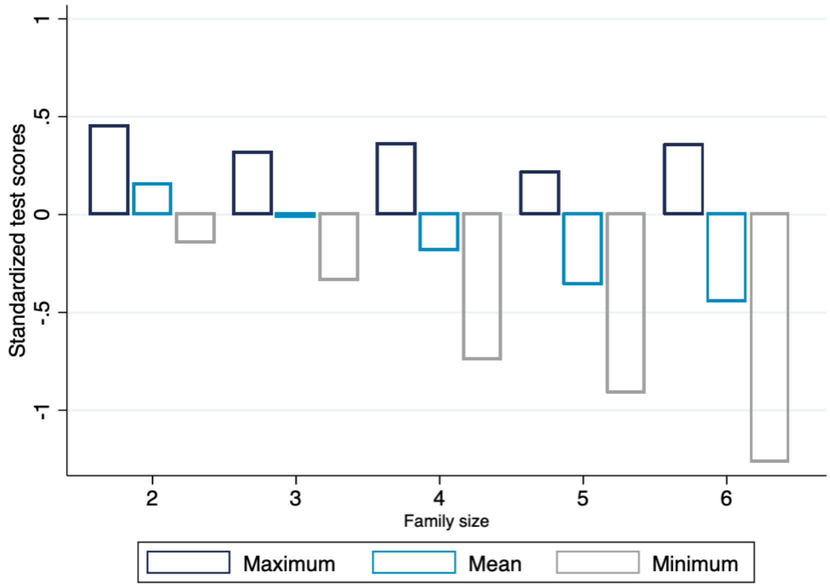
[9] Similar results for age-standardized test scores are reported in Appendix Figure 3.4.

[10] I report several robustness checks for this relation in Appendix 3.E. In particular, I show that these patterns are not specific to the Indian context, but hold more generally across a larger set of developing countries. I also show that this patterns holds is one focuses on alternative measures of child human capital, and that the relation is robust to controlling child characteristics such as gender and birth order effects. Finally, these patterns are also evident if household size is taken as endogenous, and the model is estimated by instrumental variable techniques using twin births as an instrument.

[11] Interestingly, Aizer & Cunha (2012) notice a similar pattern for a sample of poor families from the National Collaborative Perinatal Project in the US.



(A) Inequality by Family Size



(B) Human Capital Distribution by Family Size

Figure 3.2: Family Size and Inequality in Child Human Capital in India

Notes: The outcome variable is test scores. I use an age-standardized z-score, where the reference group consists of same age children children. Thus coefficients are expressed in standard deviations units. Panel A plots the within-household and between-households component of the Mean Log Deviation (MLD) measure of inequality separately by family size. By construction, the within component is zero in one-child families. Panel B plots 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. Source: India Human Development Survey (Desai et al. (2005); Desai & Vanneman (2015)).

3.3 Conceptual Framework

This section develops a simple theoretical framework to study how parents allocate resources across their children.

Preferences and constraints. Parents derive utility from their children's human capital outcome according to a Constant Elasticity of Substitution (CES) utility function that can be expressed as:

$$U(H_1 ... H_n) = (c_1 H_1^\rho + c_2 H_2^\rho + ... + c_n H_n^\rho)^{\frac{1}{\rho}} \tag{3.1}$$

where H_i is child i human capital (e.g. adult earnings), c_i are child-specific preferences (e.g. a preference for sons over daughters), and ρ regulates parental preferences for inequality in child outcomes.[12,13] This functional form assumption is standard in the literature on intra-household allocation of resources (Behrman, Pollak, & Taubman (1982); Behrman (1988)). The CES specification is very flexible as it allows a complete range of productivity-equity trade-offs. In particular, at one extreme when $\rho = 1$, the indifference curves become linear in children's outcomes as there are no inequality concerns. In this case parents act as return maximizing agents. The opposite case is the Rawlsian case when $\rho \to -\infty$; utility curves are L-shaped and parents act to equalize child outcomes. In between these two cases, parents trade-off efficiency and equity concerns.

Parents choose educational investments X_i to maximize their utility subject to two constraints. The first constraint is a budget constaint. As this is a one-period model without saving or borrowing, the budget constraint can be expressed as:

$$y = X_1 + X_2 + ... + X_n \tag{3.2}$$

where y is the total educational budget, and where the price of investments is normalized to one. We 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 share these resources between children. Notice that family size n does not have a direct effect on household resources y, but reduce the amount of per-capita resources available for investments y/n.

[12] I have made here several 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 agent. In Appendix 3.B I discuss some of the most relevant features and assumptions of the model and how some of these can be relaxed.

[13] 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. This literature suggests that parents equalize bequests across children (Behrman & Rosenzweig (2004a)).

The second constraint faced by the family is a technological constraint that maps inputs into human capital outcomes. This can be expressed as:

$$H_i = f(\theta_i; X_i; Z_i) \tag{3.3}$$

where θ_i are child initial endowments, X_i are educational investments, and Z_i are other child or family characteristics. In the empirical analysis, I consider the role of child academic ability as a measure of endowments, and parental educational expenditure as a measure of investments.[14]

Subjective beliefs. Standard models of intra-household allocations of resources rely on strong assumptions about parental knowledge of the human capital production function. In particular, these models assume that parents have perfect information about the "true" technology of skill formation in (3.3). This assumption is a very strong one, and has been shown not to hold in practice. For instance, Boneva & Rauh (2018) and Attanasio, Cunha, & Jervis (2019) show that parents hold inaccurate beliefs about the productivity of different inputs entering the production function for child human capital. To incorporate these information frictions into the model, I introduce the *perceived* human capital production function:

$$H_i = \tilde{f}(\theta_i; X_i; Z_i) \tag{3.4}$$

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 (3.3) and (3.4) play different roles in the model. The former describes the actual process of child development, while the latter represents subjective beliefs about the process, and is the relevant constraint used by parents to determine investment choices.[15]

Maximizing parental preferences subject to the constraints in (3.2) and (3.4) results in a policy function that determines parental investments. This policy function depends on household resources, preferences, on the *perceived* production function, and on endowments. Without information about the *perceived* technology, one can not derive any conclusions on parental preferences. To illustrate this point and derive closed form solutions, I assume that the actual production function is Cobb-Douglas

[14] In future work, I plan to extend the analysis to other dimensions of child endowments (e.g. physical health) and parental investments (e.g. time investments).

[15] In this paper, I do not consider the issue of how parents form these beliefs and whether these can 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 large 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 these beliefs, and whether and how these change over time should be the focus of future research.

in endowments and investments, and expressed it as:

$$H_i = \theta_i^{\alpha} X_i^{\beta} \tag{3.5}$$

where α and β are the returns to endowments and investments. Assuming that parents know the correct functional form, but that the beliefs about the productivity of different inputs can differ from actual returns, one can write:

$$H_i = \theta_i^{a} X_i^{b} \tag{3.6}$$

where a and b are the *perceived* returns to endowments and investments, and these are allowed to different from *actual* returns i.e. $a \neq \alpha$ and $b \neq \beta$.[16]

Solution. Using this parametrization of the model, one can solve for the optimal level of investments in each child (see Appendix 3.B). The optimality condition for investments in any two children in the family can be expressed as:

$$\log\left(\frac{X_i^*}{X_j^*}\right) = \frac{a\rho}{1-b\rho}\log\left(\frac{\theta_i}{\theta_j}\right) \tag{3.7}$$

Equation (3.7) shows that the interplay between preferences and the *perceived* human capital production function determines the allocation of investments across children. Without information about the perceived production function it is not possible to make any statement about parental preferences using allocation data, as observed choices are consistent with many alternative specifications of preferences and beliefs.

In particular, consider the "standard" regression used in the literature on intra-household allocation that relates parental investments to child endowments:

$$\log\left(\frac{X_i^*}{X_j^*}\right) = \gamma\log\left(\frac{\theta_i}{\theta_j}\right) = \frac{a\rho}{1-b\rho}\log\left(\frac{\theta_i}{\theta_j}\right) \tag{3.8}$$

This regression identifies a composite parameter that I label γ, which includes both preferences and beliefs. Without imposing strong assumption on such beliefs – such

[16] 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 (Attanasio, Cattan, Fitzsimons, Meghir, & Rubio-Codina (2020); Attanasio, Bernal, Giannola, & Nores (2020)). Interestingly, Attanasio, Cunha, & Jervis (2019) find that this functional form can also realistically approximate how parents perceive the production function for child human capital. Moreover, modelling child human capital as a function of investments made in one period only is arguably a simplistic assumption. A vast body of research has recently highlighted the existence of different periods of child development and has documented the existence of complementaries between early and late investments (Cunha & Heckman (2007); Johnson & Jackson (2019)). These complementaries imply that the returns to human capital investments in early periods are higher when followed by investments in later periods. For tractability, I assume that investments take place only in one initial period and leave the multi-period version of the model for future research. However, the existence of dynamic complementaries in the production of human capital could provide an incentive for parents to sustain and follow-up early investments with later ones.

that the parameters of the perceived production function correspond to the parameters of the actual technology – one can not identify parental preferences. For example, finding that investments are not sensitive to child endowments could either mean that parents care about inequality in child outcomes (ρ is close to 0), or that they believe that the returns to endowments are low (a is close to 0). In both cases the estimated γ would be close to zero. This highlights the challenge of identifying preferences from observational data. As described in the next section, in the experiment I collect direct information on parental beliefs which I combine with information on investment choices made in the experiment to identify parental preferences.

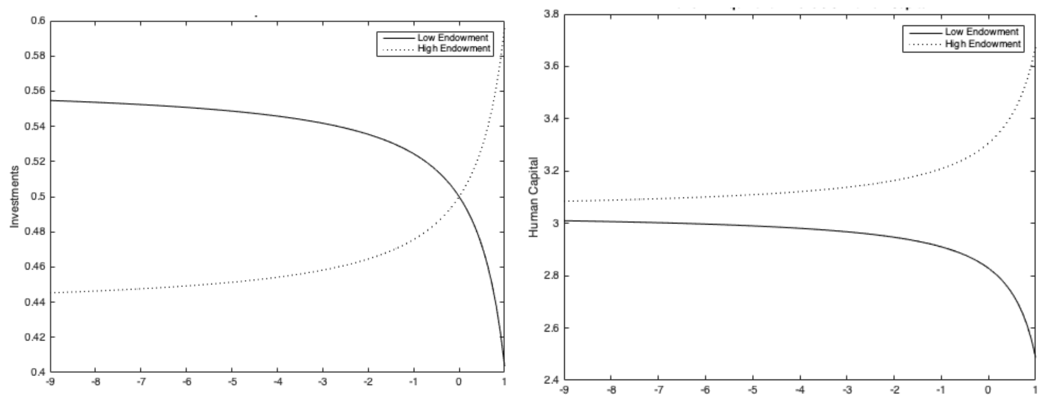
Equation (3.7) still provides us with some insights about parental investment behaviour. In particular:

- When $\gamma < 0$, parents invest more in child i whenever $\log\left(\frac{\theta_i}{\theta_j}\right) < 0$ i.e. whenever child i has an lower endowment compared to child j. In this case the parental investment strategy is *compensating.*
- When $\gamma > 0$, parents invest more in child i whenever $\log\left(\frac{\theta_i}{\theta_j}\right) > 0$ i.e. whenever child i has an higher endowment compared to child j. In this case parents adopt a *reinforcing* investment strategy.

Therefore, the model predicts that as the difference between child endowments increases, the difference in investments should increase (decrease) if parents reinforce (compensate) endowment differences. As described in the next section, I use this prediction to guide the design of the empirical strategy used in the field.

Interaction between investments and family size. So far, I have not considered how investments are affected by family size. The budget constraint shows that family size affects the amount of per- capita resources available y/n. In Appendix 3.B, I show that when reinforcement is sufficiently strong, the investments made in the human capital of the highest ability child are unaffected by family size. The intuition for this result is simple: optimizing parents invest efficiently to maximize the returns from their investment by allocating a fixed amount of resources to higher ability children, irrespectively of family size and of the endowments of other children. The model therefore predicts that investments are more reinforcing in larger families because of less per-capita resources.

Intra-household inequality. *What does the optimal investment strategy imply for intra-household inequality?* I plot the optimal investment profile in panel A of Figure 3.3. On the x-axis there is ρ and on the y-axis the optimal investments in a low and high endowment child (solid and dashed lines respectively). As ρ goes from negative to positive, investments in the lower achiever decrease while investments in the higher-achieving child increase. In panel B of Figure 3.3, I plot the equilibrium levels of human capital corresponding to the investment profile on the left panel. When ρ is



(A) Investments (B) Human capital

Figure 3.3: Preferences, Investments and the Human Capital Distribution

Notes: Panel A plots the equilibrium level of investment in child i and child j as a function of ρ. Panel B the corresponding levels of human capital as a function of ρ.

negative parents want to minimize the differences in their children's outcomes. This implies that the human capital distribution is concentrated around its mean. As ρ increases, human capital levels diverge, as parents reinforce initial differences. This figure illustrate a simple, yet key point: intra-household inequality in outcomes is only consistent with the case where parents are not averse to intra-household inequality i.e. $\rho > 0$.

How does family size contribute to intra-household inequality? As discussed above, when investments reinforce child endowments the level of resources invested in the highest achieving child will be unaffected by family size. The implications of this investment behaviour – coupled with the existence of decreasing returns to investments – is that as family size increases there is a shallow gradient in maximum child quality and a steep drop in the quality of the lowest achieving child.[17] This implies that the level of intra-household inequality will increase with family size. In particular, low endowment children are penalized by increases in family size both because by having more siblings they face more competition over resources, and because they are likely to fare poorly compared to their siblings when competing with them.

3.4 Lab-in-the-field Experiment

The conceptual framework highlights the different role of preferences, beliefs and household resources in determining investment choices and intra-household inequality in child outcomes. It also illustrates the challenges that observational data pose for the identification of key parameters and mechanisms of interest. The reason for this challenge is a twofold identification issue. First, realized choices may be consistent

[17]Decreasing marginal product of investments means that a reduction in investments will have a much larger impact on child human capital when it happens at a low baseline level of investments, that when it happens at a high baseline level. Appendix 3.B and Appendix Figure 3.9 discuss in greater detail the intuition behind this result.

with many alternative specifications of preferences and beliefs (Manski (2004)). Second, resource constraints might prevent parents from selecting their preferred choices, breaking the direct connection between observed outcomes and parental preferences (Baland & Robinson (2000)).

To overcome these identification challenges, I design and implement a lab-in-the-field experiment with parents of primary school children in India. In the experiment, I use a novel survey methodology based on hypothetical scenarios that is closely guided by the theoretical model. This allows me to identify parental beliefs about the human capital production function. I then use hypothetical scenarios, to collect stated investment choices. Combining the choices in the second stage with the beliefs from the first stage, I identify parental preferences for intra-household inequality. Furthermore, as scenarios vary in terms of resources available to the family, I can study whether household resources are important in determining investment decisions. Next, I describe in detail the measurement tools, the experimental procedures and how I combine experimental measures to identify primitive parameters of interest.

3.4.1 Parental Beliefs

Measurement. To elicit parental beliefs about the human capital production function, I build on the work by Cunha, Elo, & Culhane (2013), Boneva & Rauh (2018) and Attanasio, Boneva, & Rauh (2019) and use hypothetical scenarios. This strategy has the advantage that one can vary one input at the time while holding *all* other inputs fixed, thus identifying the *perceived* productivity of different inputs.

In the experiment, I presented each respondent with a series of hypothetical stories about a representative family. Guided by the theoretical framework, I focus on the role of perceived returns to child endowments and parental investments, and on their perceived complementarity or substitutability, as these are the key parameters that matter for the allocative decision. In each scenario, I exogenously varied one input and asked the respondent to report what they believed the future earnings of the child would be at age 30 (this corresponds to H_i in the theoretical framework). As a robustness check, I also asked parents to state what they believed the educational attainment of the child would be in each hypothetical scenario.

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 an hypothetical child rather than their own, because this allowed to vary only one input at the time. One particularly important input in this context is child academic ability. Exogenous variation in this input would clearly not have been possible if I asked respondents about their own child.

The experimental procedures worked as follows. Surveyors asked respondents to think about a representative family that lives in a neighbourhood like their own.

The family has *two* children who attend the same school and are identical in many respects.[18] While the first child (Child H) has an high academic ability, the second child (Child L) has a low academic ability (this corresponds to θ_i in the model). Specifically, while Child H was described as being *among the top three students in his/her class*, Child L was described as being *among the bottom three students in his/her class.*[19] Scenarios then varied in term of the amount of monetary investment made by the family in each child. After presenting the scenario, surveyors asked respondents to report what they believed the outcome would be for each child. The respondent's answer was recorded, and the experiment moved to the next scenario. To insure understanding, all scenarios were presented to the respondent with the help of a visual aid that sketched the main features (see Appendix Figure 3.5 for an example of the visual aids used in the field).[20]

Finally, to understand whether parents perceived these returns to differ by gender, I randomized respondents in two groups so that one group saw two boys, while the other group saw one boy and one girl.

Identifying the perceived production function. Comparing responses across scenarios and between children one can identify: (i) the perceived returns to investments, (ii) the perceived returns to endowments, and (iii) the perceived complementarity or substitutability between these two inputs. For example, by comparing responses across scenarios where investments are high to the corresponding scenarios where investments are low one can indentify 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):

$$y_{i,j,k} = \alpha_0 + \alpha_1\theta_{j,k} + \alpha_2 I_{j,k} + \alpha_3\theta_{j,k} \times I_{j,k} + \gamma_i + u_{i,j,k} \tag{3.9}$$

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, $\theta_{j,k}$ is a dummy equal to one if child k' ability is high, $I_{j,k}$ is a dummy equal to one if investments in child k are high, and γ_i are respondent fixed effects. The coefficients α_1 and α_2 identify the perceived returns to endowments and investments, while the coefficient α_3 identifies the perceived complementarity ($\alpha_3 > 0$) or substitutability ($\alpha_3 < 0$) between these two inputs.

[18] Presenting respondents with an hypothetical family with two children rather than two distinct families with one child each has the advantage of holding fixed many of the unobserved factors that might matter for child outcomes and that vary between families (e.g. parental income and the family environment).

[19] As it is common in developing countries, parents in India regard school performance as the most important measure of academic ability. This is also the first reliable and objective measure of academic ability that parents have access to.

[20] Appendix 3.F presents the exact wording of some relevant questions used in the survey.

3.4.2 Investment Choices

Measurement. Having characterised the features of the perceived human capital production function, in the second round of the experiment I collect stated investment choices. The measurement tool used in this stage is similar to the one used to elicit parental beliefs. 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, I ask them to select their favourite allocation choice.[21]

Respondents are presented with a representative family who makes a decision about how to distribute educational *expenditure* across their children. Guided by the theoretical model, across scenarios I exogenously varied: (i) children's endowments, and (ii) the amount of resources the family can spend on their children's education. In particular, the model predicts that as the difference between childrens' endowments increases, the difference in investments should increase (decrease) if parents reinforce (compensate). Importantly for identification, I therefore varied the *difference* in endowments 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 the scenario, respondents are asked how they would allocate resources. Specifically, they are asked to distribute some tokens to reflect 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 parental 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.[22]

Comparing parental allocations across scenarios, I study whether investments reinforce or compensate endowments' differences. To understand whether choices are affected by household resources, I then compare responses in scenarios where resources are high with the choices in the corresponding scenario with low resources. I decided to vary the resources available to the hypothetical family rather than the number of children in the family (as in the theoretical framework) to avoid respondents' confusion and for consistency with the scenarios used to elicit beliefs. Moreover, as discussed later, most families in the sample have only two children, so that the hypothetical scenarios are particularly relevant and realistic for them. Importantly, in the model

[21] 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 & Andrew (2019); Ameriks, Briggs, Caplin, Shapiro, & Tonetti (2020); Ameriks, Briggs, Caplin, Lee, et al. (2020)).

[22] The exact wording of some relevant questions used in the survey is presented in Appendix 3.F.

the only way in which family size affects human capital investments is through per-capita resources. Finally, to understand whether investments differ by child gender, I randomized the gender of one of the two children across respondents so that one group saw two boys, while the other group saw one boy and one girl.

Characterizing the investment strategy. To study parental investments, I estimate variants of the following empirical specification using ordinary least squares:

$$s_{i,j} = \beta_0 + \beta_1 diff_j + \gamma_i + u_{i,j} \tag{3.10}$$

where i indicates the respondent and j the scenario, $s_{i,j}$ is the share of total resources allocated to the higher endowment child, and $diff_j$ is a dummy variable that is equal to 1 if in scenario j the difference between the two children is large and zero otherwise. The sign of β_1 pins down whether parental investments are reinforcing ($\beta_1 > 0$) or compensating ($\beta_1 < 0$). To test whether resources matter to explain allocations, I expand equation (3.10) and estimate:

$$s_{i,j} = \beta_0 + \beta_1 diff_j + \beta_2 res_j + \beta_3 diff_j \times res_j + \gamma_i + u_{i,j} \tag{3.11}$$

where res_j is a dummy variable that takes value 1 if in scenario j resources are high. The sign of β_3 identifies if reinforcement is stronger when resources are lower ($\beta_3 < 0$).

3.4.3 Combining Measures to Identify Preferences

While estimates of equations (3.10) and (3.11) are informative of whether parents reinforce or compensate, without further assumptions on parental beliefs we are not able to say anything about preferences for intra-household inequality. This point can be easily illustrated by looking again at equation (3.8), which is reported here for convenience:

$$\log\left(\frac{X_i^*}{X_j^*}\right) = \gamma \log\left(\frac{\theta_i}{\theta_j}\right) = \frac{a\rho}{1-b\rho}\log\left(\frac{\theta_i}{\theta_j}\right)$$

A regression of parental allocations on child endowments identifies a composite parameter that includes both preferences and beliefs. Without imposing strong assumption on such beliefs – such that they coincide with the parameters of the production function – one can not identify preferences. For example, if one were to find that parental allocations are insensitive to endowments, it might either be that parents have concerns for intra-household inequality, or it could be that parents perceive the returns to endowments to be particularly low. In both cases the estimated γ would be close to zero.

Combining the experimental data on beliefs and choices, I can identify parental preferences for intra-household inequality. The intuition for the identification result is simple. A regression of expected child outcomes on investments and endowments identifies the parameters of the perceived production function (a and b). Moreover, choices in the intra-household allocation module can be used to identify the reduced form parameter γ. Armed with these parameters one can identify parental preferences as follows:

$$\rho = \frac{1}{a} \times \left[\frac{1}{\gamma} + \frac{b}{a}\right]^{-1} \tag{3.12}$$

A consistent estimator for ρ can then be obtained by replacing the parameters in (3.12) with the corresponding OLS estimates from equations (3.9) and (3.10).[23,24]

Stated and revealed preferences. One natural question is whether preferences recovered from data on hypothetical choices relate to observed behaviour. To address this question, I also collect data on actual investments made by parents in the form of child specific educational expenditure. I investigate whether respondents predicted to be less inequality averse in the experiment systematically make more unequal choices when it comes to distribute actual resources across their 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.

3.5 Data and Descriptive Statistics

The experimental sample consists of 504 households with children in the urban slums of Cuttack, Odisha, India. 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.[25,26] In 2013, a sample of poor women with young children (aged 10 to 20 months then) was identified through a door-to-door census. In 2019, we aimed at re-interviewing all households in the original sample.[27] Survey respondents

[23]Given consistency of the OLS estimator, using the continuous mapping theorem and Slutsky theorem one can prove consistency of the estimator for ρ. Standard errors and confidence intervals can be obtained using bootstrap methods.

[24]Notice that the identification strategy can also account for to the possibility that parents might hold inaccurate beliefs about their children's endowments, as in Dizon-Ross (2019). As I have control over *all* the characteristics of the hypothetical scenario, I can precisely describe child endowments to the respondent, thus avoiding the issues related to the fact that objective measures of child endowments might not accurately reflect parental beliefs.

[25]See Andrew et al. (2019) for details of the intervention and its short term impacts on child human capital.

[26]Appendix Tables 3.5 and 3.6 show that the treatment effects on parental preferences and beliefs are negligible. This is consistent with the results from the first follow-up, showing 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 the following sections.

[27]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).

Table 3.1: Summary Statistics

	Mean	S.D
A. Household characteristics		
Primary caregiver has no formal education	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 the household statistics, Panel B the statistics 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.

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, whenever possible, in a quiet and private environment.

Table 3.1 reports the summary statistics for the sample. It shows that this is an economically and socially disadvantaged population: only 50 percent of children's primary caregivers have any formal education, and just over 30 percent of household are attached to the sewage system. 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 everyday. The table also shows that the percentage of sons among children is 48 percent, which suggests that, at least in term of this indicator, there is no evidence of a strong son preference in this context (if anything the sex ratio is slightly skewed towards females, and this is also true if we consider all household members and not children specifically).[28]

[28]In 2019, the national sex ratio was 108 boys per 100 girls (United Nations (2019)). The natural

3.6 Results

This section is organized as follows. Section 3.6.1 presents the results on parental subjective beliefs about the human capital production function. The results on stated choices and what they imply for parental preferences are presented in section 3.6.2. In section 3.6.3 I relate elicited preferences to actual educational investments made by the respondents in their own children.

3.6.1 Beliefs

I present the estimates of equation (3.9) in Table 3.2. I start by regressing perceived (log) earnings on a dummy for high endowment and a dummy for high investment in column 1. I subsequently control for child gender and for the interaction between endowment 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 by parents).[29] A few interesting findings emerge.

First, parents perceive the returns to initial ability to be large: high endowments are associated with an increase in earnings of between 70 percent to 80 percent. At the sample mean of expected earnings this corresponds to an increase of roughly INR 13,000 to 14,000. Interestingly, a mediation analysis suggests that almost 50 percent of the effect of child endowment comes through higher educational attainment (column 5). In particular, as show in Appendix Table 3.7, parents believe that higher ability children will achieve two more years of schooling compared to lower ability children. In turn, one year of schooling increases earnings by 16.9 percent. Second, column 2 shows that investments are perceived to increase earnings by 24.6 percent. Third, the results in column 3 imply that endowments and investments are perceived as complements: parents believe that the returns to investments are 10.3 percent higher for a high endowment child. This perceived complementarity provides a potential rationale for parents to invest more in higher ability children.

Benchmarking perceived returns. Table 3.2 reports the coefficients associated with a binary increase in the relevant input (i.e. a change from a *low* level of the input to an *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 endowments 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 my sample, parents expect boys to earn

sex ratio at birth is 105 boys per 100 girls.

[29] Results for educational attainment follow a quantitative similar pattern and are presented in Appendix Table 3.7.

Table 3.2: Perceived Production Function

Outcome variable: Log earnings at age 30	(1)	(2)	(3)	(4)	(5)
High endowment	0.768*	0.848*	0.796*	0.717*	0.371*
	(0.021)	(0.030)	(0.031)	(0.022)	(0.028)
High investment	0.246*	0.246*	0.194*	0.194*	0.152*
	(0.009)	(0.009)	(0.011)	(0.011)	(0.010)
Boy		0.159*	0.159*	-	-
		(0.041)	(0.041)		
High endowment x High investment			0.103*	0.103*	0.102*
			(0.013)	(0.013)	(0.013)
Belief about child education					0.169*
					(0.012)
Family fixed effects				✓	✓
Observations	7920	7920	7920	7920	7920

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 error clustered at the family level are reported in brackets. *High endowment* 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, 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 will achieve in scenario j. *denotes 5% significance.

on average 16 percent more than girls at age 30 (columns 2 and 3 of Table 3.2). This figure is quite close to the actual gender-gap in urban workers' earnings of 22 percent (ILO (2018)). Interestingly, I also find that while parents believe that girls on average will command less resources than boys as adults, they do not perceive the returns to endowments or investments to substantially differ by gender (see Appendix Table 3.8). 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 treatment of women and men in the labour market.[30]

Beliefs heterogeneity. The estimates in Table 3.2 represent *average* beliefs about the returns to different inputs. To uncover potential heterogeneity in these beliefs, I follow Attanasio, Boneva, & Rauh (2019) and construct an individual measure of perceived returns. For example for investments, I compute the difference between respondent's expected earnings in the scenarios in which investments are high and scenarios in which they are low and take an average across scenarios. I plot the results in Appendix Figure 3.6. Panel A displays the empirical cumulative distributions of individual perceived returns to investments. The figure reveals a substantial degree of heterogeneity in perceived returns and that, consistently with the findings form Table 3.2, perceived returns are higher for children higher ability children. By comparing expected earnings across high and low ability children while holding investment

[30]Essentially, we can interpret this as a shift in the total factor productivity in the production function, while leaving unchanged the other technology parameters.

Table 3.3: Intra-household Allocation of Resources

Outcome variable: share of resource to child H	(1)	(2)	(3)	(4)
Difference in endowments	0.078*	0.078*	0.102*	0.102*
	(0.005)	(0.006)	(0.008)	(0.009)
High resources			0.028*	0.028*
			(0.007)	(0.008)
Difference in endowments × High resources			-0.048*	-0.048*
			(0.008)	(0.009)
Boy	-0.001		-0.002	-
	(0.008)		(0.008)	
Family fixed effects		✓		✓
Observations	1980	1980	1980	1980

Notes: The outcome variable is the share of total resources invested in child H. This variable ranges from 0 to 1. The average of this variable is 0.52. Columns 1 and 3 display the OLS results, while columns 3 and 4 further includes family fixed effects. Robust standard error clustered at the family level are reported in brackets. *Difference in endowments* is a dummy variable that takes value 1 if in scenario j the difference between the two children's endowments 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 5% significance.

fixed, I also compute the individual perceived returns to child endowments. The distribution of these perceived returns is shown in panel B, and also shows substantial heterogeneity.

3.6.2 Investment Choices and Preferences

Table 3.3 reports the estimates of equations (3.10) and (3.11). 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 endowments increases parents reallocate resources within the family 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 Child H, which at the "average allocation" corresponds to a 15 percent increase. The positive coefficient implies that parental investments are reinforcing.

Interestingly, I do not find evidence that parental choices depend on the gender of the child. I also tried estimating equations (3.10) and (3.11) separately for the two different subgroups (defined based on the gender of the two children) and found very similar results. Although the previous literature does not always find evidence of differential treatment of boys and girls, recent work on India has shown 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 3.3, one has to keep in mind that the input being allocated here is educational *expenditure*. Consistently with my findings, previous research has found no evidence of parents spending differently on boys and girls (Deaton (1989, 1997)). Moreover, in interpreting these results one has to consider that in urban Odisha there is little evidence of girls receiving less human capital investments compared to boys, nor there is evidence of a skewed sex-ratio in the sample (Table 3.1), suggesting that son preferences might be less prevalent in the context of this study.[31]

It might then seem odd that parents equally allocate educational investments between boys and girls, despite them perceiving girls to be able to command less resources as adults (Table 3.2). One potential reasons 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 & Andrew (2019); Ashraf, Bau, Nunn, & Voena (2020)).[32]

Turning to the role of household resources, 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 endowments 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 endowment difference is large compare to when it is small. This gap in investments is halved when resources are high. This result highlights the critical role that financial constraints have in determining investments in child human capital. Importantly, the finding implies that resources are key to explain investment behaviour within the household, and complements previous research showing that resources are an important determinant of investment differences between families (Attanasio & Kaufmann (2009); Lochner & Monge-Naranjo (2012)). Therefore, it seems plausible that relaxing credit constraints can contribute to close investments gaps across children, potentially also resulting in lower inequality in outcomes.

Preferences. As discussed earlier, a regressions of investments on endowments only identifies the composite parameter γ. This parameter comprises both parental preferences for inequality and their perceptions about the production function. Using parental beliefs from the previous section, I identify preferences using the procedure outlined above. I find that the implied ρ is positive and statistically significant at the 95% significance level, implying that in this setting investment choices are to

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

[32] This interpretation is also consistent with the results in Appendix Table 3.7 which shows that parents do not expect girls and boys' educational attainment to differ substantially.

some extent driven by efficiency considerations (the point estimate is 0.449, with an associated standard error of 0.041).

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 3.7. 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 ρ.[33]

3.6.3 Stated and Revealed Preferences

One 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 rich information on current investments made by the respondents in their own children, and relate child-specific investments to child endowments. To measure endowments 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. Importantly, what matters to understand intra-household allocations is whether these beliefs (more precisely the difference in beliefs between two children) explains parental investments.[34]

The results are presented in Table 3.4. I start by running the "standard" regression that relates the difference in educational expenditure between two children to the difference in their endowments, controlling for other observable characteristics including child age and gender. The results in column 1 suggest a positive and significant correlation between child ability and investments. In particular, the point estimate implies that a 10 percent increase in the difference between children's endowments is associated with an increase in educational expenditure gap of INR 378 in favour of the higher ability child. This corresponds to 5.6 percent of total yearly educational expenditure. In Appendix Table 3.9, I further show that higher abillity children are more likely to attend a private school and less likely to work, suggesting that these decisions expand beyond educational expenditure at a particular point in time, with potentially important longer-term effects.

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

[34] As a robustness check, I also experimented using two more "objective" measures of child endowments: whether the child suffered from any health condition in the first three years after birth and the health status of the child as reported by the primary caregiver (results available upon request). Both measures implied qualitatively similar results to those reported in Table 3.4.

Table 3.4: Actual Parental Investments

Outcome variable: difference in expenditure	All (1)	Low ρ (2)	High ρ (3)
Difference in endowments	378.044*	221.317	470.211*
	(88.320)	(132.537)	(117.972)
Mean expenditure	6662	6662	6662
Observations	1100	552	548

Notes: The outcome variable is the difference in investment between two children, as measured by educational expenditure in Rupees. Column 1 report the results in the full sample, while columns 2 and 3 report separate results for two separate sub-samples as defined by their inequality aversion (low ρ means higher inequality aversion). These two groups are defined based on whether the estimated ρ falls above or below the sample median. Robust standard error clustered at the family level are reported in brackets. Controls include child age and gender. *denotes 5% significance.

I next turn to the more important 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 median). Column 2 and 3 report the results from this exercise. 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 twice as large as that in column 2 and statistically different from zero. This estimate implies that a 10 percent increase in the difference in endowments is associated with an increase in the educational expenditure gap between children of INR 470. This corresponds to 7 percent of the yearly educational expenditure. On the other hand, for inequality averse families this figure is only 3.3 percent of total yearly educational expenditure.

The fact that my experimentally elicited measure of parental preferences maps into actual investment behaviour is reassuring, as it adds credibility to the research design and to the use of hypothetical scenarios to identify primitive parameters of interest.

3.7 Conclusions

This paper studies the role of parental investments as a determinant of intra-household inequality in child human capital. I first document that across developing countries within household variation explains between 30 to 50 percent of overall inequality in child educational attainment. By looking at the human capital distribution within a 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 informative about parental investment behaviour, and consistent with the differential treatment of children in terms of educational investments.

In order to understand how parents make these investment decisions, I combine a theoretical model of household behaviour with experimental data from India. The model highlights that investments depend on parental preferences, beliefs and household resources. To mitigate the identification problem posed by observational data, I design a lab-in-the-field experiment with parents of primary school children. I use a novel survey methodology based on hypothetical scenarios to collect data on subjective expectations and stated choices with and without financial constraints.

Several key results emerge from this study. First, I find that parents perceive endowments and investments to be highly productive, and to be complements in the production of human capital. This suggests that parents should invest more in higher ability children 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 endowment differences. This suggest that in this setting investment choices are to some extent driven by efficiency considerations. Third, I show that financial constraints are important in explaining household investments, as parents reinforce more strongly when per-capita resources are lower. Finally, I show that experimentally elicited preferences relate to *actual* household behaviour, and that respondents who are identified as less inequality averse in the experiment, systematically make more unequal allocations.

The results in this paper indicate that early levels of human capital have a key role in driving inequality *within* the family, complementing previous research looking at their role in explaining inequality *between* families. Children with low initial levels of human capital are particularly penalized in two ways. First, through a biological channel coming from the human capital productions function. Second, they are penalized by parental investments which reinforce these initial differences.

These findings have important implications for policy. First, the fact that parents respond to child human capital levels suggest that early childhood interventions can generate large *direct* positive effects on human capital outcomes, as reported for example by Heckman (2006), but also that they have the potential to produce important *indirect* effects through parental endogenous investment responses, thus magnifying total returns. These results are also important for intra-household inequality, as they imply that acting on human capital in the early years can affect the way parents distribute educational resources across their children.

The findings also point to the important role that household resources have to explain human capital outcomes. They suggest that relaxing resource constraints could lead to more equal allocations between children, resulting in improvements in child development, particularly for lower ability children. However, for such intervention to be effective one needs to understand how parents would distribute those additional resources across their children. One commonly used policy to improve child outcomes in developing countries are cash transfers to families (Fiszbein & Schady (2009)). However, giving cash to a family might not be sufficient to improve the outcomes of *all* children, as this will crucially depend on the allocation of resources within the household. For instance, Barrera-Osorio, Bertrand, Linden, & Perez-Calle (2011) report that parents adjust their investments in response to a conditional cash transfer program in Colombia by diverting educational resources away from *non-target* children towards *target* siblings. This result is consistent with the fact that the intervention might have made more salient to the parents the returns to investing in one specific child in the family.

More generally, the results in this paper suggest that the effects of interventions aimed at improving children's outcomes will crucially depend on parental endogenous responses, which are mediated by their preferences, beliefs and constraints. Understanding how to incorporate these endogenous responses in the design of *effective* policies should be an important area of future research.

Appendix 3.A Mean Log Deviation Measure of Inequality

Figure 3.1 and Figure 3.2 use the Mean Log Deviation Measure of Inequality (MLD) to decompose overall inequality in child human capital outcomes in a within-household and between-households components. The MLD can be expressed as:

$$MLD = \frac{1}{N} \sum_i ln \frac{\bar{y}}{y_i} \tag{3.13}$$

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

$$MLD = \sum_j \frac{N_j}{N} MLD_j + \sum_j \frac{N_j}{N} ln \frac{\bar{y}}{\bar{y}_j} \tag{3.14}$$

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).

Appendix 3.B Model Appendix

Close form solution for investments

In this section, I derive a closed form solution for investments. Maximizing (3.1) subject to (3.2) and (3.6) one can get to the following closed form solution for investments in child i:

$$X_i^* = y \frac{\theta_i^{\frac{a\rho}{1-b\rho}}}{\sum_{j=1}^{n} \theta_j^{\frac{a\rho}{1-b\rho}}} \tag{3.15}$$

Computing the ration of X_i^* to X_j^* and taking the log we get equation (3.7).

Human capital investments and family size

For sufficiently high values of γ, the child with the highest endowment will receive roughly the same share of household resources independently of family size. I define the largest endowment in the family as θ_{max}. For any family size n, the model implies that educational investments in the highest endowment child are:

$$X^*_{max}(n) = y \frac{\theta_{max}^{\frac{a\rho}{1-b\rho}}}{\underbrace{\theta_{min}^{\frac{a\rho}{1-b\rho}} + ... + \theta_{max}^{\frac{a\rho}{1-b\rho}}}_{\text{n terms}}} = y \frac{\theta^{\gamma}_{max}}{\theta^{\gamma}_{min} + ... + \theta^{\gamma}_{max}} \tag{3.16}$$

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 (3.16). As $\gamma \rightarrow \infty$:

$$\lim_{\gamma \rightarrow \infty} X^*_{max}(n) = \lim_{\gamma \rightarrow \infty} y \frac{1}{\left(\frac{\theta_{min}}{\theta_{max}}\right)^{\gamma} + ... + 1} = y \tag{3.17}$$

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 $X^*_{max}(n) \rightarrow X^*_{max}(n+1)$.

Intuition Figure 3.9 provides the intuition behind this mechanisms. The top panels plot ρ on the x-axis (holding fixed a and b) and the share of total resources allocated to an high endowment child (left panel) and a low endowment child (right panel) on the y-axis. In each plot there are two lines: the blue line represents the share of resources the child receives in a family with n children, while the black line represents the share she receives in a family with $n+1$ children. The graph shows that as ρ the resources devoted to θ_H increase, while the resources allocated to θ_L decreases. Interestingly, as ρ increases the shares of resources allocated to each child in families of different sizes converge.

The bottom panels plot the *change* in resources as we move from a family with n to $n+1$ children. This corresponds to the vertical distance between two lines in the corresponding top panel. The figures show that as family size increases low endowment children are more penalized in terms of resources, because of more competition implied by a larger family size. This result, coupled with the existence of decreasing returns to investments, explains why as family size increases there is a shallow gradient in maximum child quality and a steep drop in the quality of the lowest achieving child.[35]

Discussion of the model

I this section, I discuss some of the most relevant features and assumptions of the model and how some of these assumptions can be relaxed. I conclude with a discussion of the model's implications for the Quantity-Quality trade-off.

[35]Decreasing marginal product of investments means that a reduction in investments will have a much larger impact on child human capital when it happens at a low baseline level of investments, that when it happens at a high baseline level.

Gender preferences. By including weights to child human capital in the parental utility function (3.1), the model is general enough to incorporate social norms such as gender preferences or other details that are important in specific contexts. For instance, a literature suggests that in the Indian context parents have 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)).

In solving the model, I set these utility weight equal to each other. This is because I do not find strong evidence of a gender preference in the reality of my setting. In particular, there is no evidence of a skewed sex ratio in the sample (see Table 3.1). This finding is consistent with the fact that the state of Odisha – where the data used in this paper come from – has one of the less skewed sex ratios in the country (see Appendix Figure 3.8).[36] Moreover, previous work demonstrates a non-significant discrimination against girls in Odisha in terms of post-natal investments (IIPS (2001); IIPS (2008); Padhi (2001)).[37]

Child endowments. One important feature of the model is child birth endowment. This paragraph discusses what these endowments are and whether they are observed by the parents.

Endowments at birth encompass a variety of different characteristics that include both health and cognitive ability. Several recent studies that rely on measures of child health at birth to proxy for endowments often have physical health in mind as the key dimension. Empirically this is operationalised using birth weight as a measure of human capital at birth (Datar, Kilburn, & Loughran (2010); Hsin (2012)). Theoretically this assumption is not needed and is often made for tractability, given that obtaining information on child cognitive endowments can be even more challenging. One exception is Adhvaryu & Nyshadham (2016) that considers child cognitive abilities by exploiting a large-scale iodine supplementation program in Tanzania. Therefore, in practice endowments at birth might comprise a bundle of health and cognitive skills. As illustrated by the model, what matters to study parental behaviour is that these endowments affect the return to investing in a child and are thus relevant to determine child long run human capital outcomes. In the empirical analysis, I consider

[36] Chao, Guilmoto, KC, & Ombao (2020) predict that by 2030 the sex ratio in Odisha (male to female) will be the third lowest in the country (following Kerala and Chhattisgarh) and will be of 105 males per 100 females, which corresponds to the natural sex ratio at birth WHO (2019).

[37] In terms of educational investments, in urban areas school attendance is the same for boys and girls in the age groups 6-10 years and 15-17 years, and slightly higher for girls than boys in the age group 11-14 years (81% of girls compared with 78% of boys). Similarly, in terms of health investments and outcomes, the infant and under-five mortality rates are 23-26 percent higher for boys than for girls. Among surviving children, girls and boys are about equally likely to be undernourished. Girls are also more likely than boys to be fully vaccinated (55% of girls, compared with 49% of boys). Table 3.10 further shows that compared to the rest of the country, girls in Odisha are less likely to belong to larger families.

child cognitive ability as a measure of endowments.

This discussion leads to the question of whether parents can observe endowments. While this assumption is implicitly made by most models relating endowments to subsequent parental behaviour, there is no explicit discussion of whether it holds in practice. A large medical literature suggest that parents are indeed able to recognise their child's endowment from very early ages. For instance, Channon (2011) indicate that mother's perception of their child's size is a good proxy for actual birth weight. On the other hand, recent evidence from the economics literature suggests that sometime parents might have inaccurate perceptions about their children's cognitive ability, causing the mis-allocation of education investments (Dizon-Ross (2019)).[38] Importantly, the theoretical framework can be easily extended to allow for this possibility. Specifically, by replacing *actual* endowments in the perceived human capital production function with *perceived* endowments one can derive similar implications for parental investment behaviour. What will matter now to determine the allocation of resource across siblings is *perceived* child endowments. Moreover, the empirical strategy that I use, which relies on the use of hypothetical scenarios, is robust to the possibility that parents might have inaccurate beliefs about endowments.

Fertility choices. One assumption made 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 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 endowments so that parents are more likely to increase fertility after giving birth to a low endowment child.[39] 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,

[38] The psychology literature suggests that mothers are able to assess and react to signals of cognition in their infant children from as early as a few days after birth (Brazelton (1984); Bullowa (1979); Susman-Stillman, Kalkoske, Egeland, & Waldman (1996)).

[39] 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)).

children born with low initial conditions 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 endowments 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 the endowments of their children, 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 explain why the empirical findings in the Quantity-Quality literature are mixed, with some studies finding evidence in favour of a trade-off (Rosenzweig & Wolpin (1980); Hanushek (1992); Rosenzweig & Zhang (2009); Mogstad & Wiswall (2016)), while other against (Black, Devereux, & Salvanes (2005); Angrist, Lavy, & Schlosser (2010); Cáceres-Delpiano (2006)). What the model suggests is that family size *per se* might have little effect on child human capital, what matters for child outcomes is the effect that family size has on per-capita resources, combined with parental investment decisions.

Appendix 3.C Appendix Tables

Table 3.5: Effect of RCT Treatment Status on Perceived Production Function

Outcome varibale: Log earnings at age 30	(1)	(2)
High endowment	0.860*	0.808*
	(0.056)	(0.056)
High investment	0.221*	0.169*
	(0.014)	(0.018)
Treatment	0.052	0.046
	(0.068)	(0.068)
High endowment × Treatment	-0.021	-0.009
	(0.058)	(0.060)
High Investment × Treatment	0.026	0.038
	(0.020)	(0.026)
High endowment × High investment	-	0.103*
		(0.019)
High endowment × High Investment × Treatment	-	-0.025
		(0.031)
Boy	0.147*	0.147*
	(0.053)	(0.053)
Observations	4960	4960

Notes: This table presents analogous coefficients and standard errors to those presented in Table 3.2 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 3.2 , and for column 2 is column 3 of Table 3.2. Robust standard error clustered at the family level are reported in brackets. * denotes 5% significance.

Table 3.6: Effect of RCT Treatment Status on Allocation of Resources

	Share of resources to child H	
	(1)	(2)
Difference in endowments	0.077*	0.109*
	(0.010)	(0.015)
Treatment	-0.004	-0.006
	(0.013)	(0.019)
Difference in endowments × Treatment	0.001	-0.014
	(0.014)	(0.021)
High resources	-	0.033*
		(0.011)
Difference in endowments × High resources	-	-0.064*
		(0.014)
High resources × Treatment	-	0.005
		(0.016)
Difference in endowments × High resources × Treatment		0.032
		(0.021)
Boy	0.006	0.006
	(0.011)	(0.011)
Observations	1980	1980

Notes: This Table presents analogous coefficients and standard errors to those presented in Table 3.3 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 3.3, and for column 2 is column 3 of Table 3.3 . Robust standard error clustered at the family level are reported in brackets. * denotes 5% significance.

Table 3.7: Perceived Production Function (Educational Attainment)

Outcome varibale: Educational attainment (in years)	(1)	(2)	(3)	(4)
High endowment	2.042*	2.043*	2.039*	2.038*
	(0.046)	(0.056)	(0.056)	(0.048)
High investment	0.252*	0.252*	0.248*	0.248*
	(0.016)	(0.016)	(0.020)	(0.021)
Boy		0.002	0.002	-
		(0.062)	(0.062)	
High endowment x High investment			0.008	0.008
			(0.024)	(0.025)
Family fixed effects				✓
Observations	7920	7920	7920	7920

Notes: This table reports parent perceived returns. 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 error clustered at the family level are reported in brackets. * denotes 5% significance.

Table 3.8: Perceived Production Function by Gender

Outcome varibale: Log earnings at age 30	Girls			Boys		
	(1)	(2)	(3)	(4)	(5)	(6)
High endowment	0.778*	0.729*	0.340*	0.758*	0.704*	0.401*
	(0.035)	(0.037)	(0.043)	(0.023)	(0.024)	(0.036)
High investment	0.250*	0.202*	0.157*	0.242*	0.187*	0.148*
	(0.012)	(0.015)	(0.014)	(0.013)	(0.016)	(0.015)
Hight endowment × High investment	-	0.097*	0.095*	-	0.109*	0.109*
		(0.019)	(0.018)		(0.018)	(0.018)
Belief about child education	-	-	0.189*	-	-	0.150*
			(0.017)			(0.018)
Family fixed effects		✓	✓		✓	✓
Observations	3968	3968	3968	3952	3952	3952

Notes: The table report coefficients analogous to those presented in Table 3.2 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 error clustered at the family level are reported in brackets. High endowment 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 5% significance.

Table 3.9: Actual Parental Investments: Additional Outcomes

Outcome variable:	Attends private school			Child works		
	All	Low ρ	High ρ	All	Low ρ	High ρ
	(1)	(2)	(3)	(4)	(5)	(6)
Difference in endowments	0.111*	0.111*	0.110*	-0.058*	-0.046*	-0.073*
	(0.011)	(0.016)	(0.016)	(0.005)	(0.007)	(0.006)
Mean outcome		0.201			0.149	
Observations	995	496	499	711	353	358

Notes: The outcome variable is the difference in investment between two children, in the outcome variable shown in the column header. Column 1 and 4 report the results in the full sample, while columns 2, 3, 5 and 6 report separate results for two separate sub-samples as defined by their inequality aversion (low ρ means higher inequality aversion). These two groups are defined based on whether the estimated ρ falls above or below the sample median. Robust standard error clustered at the family level are reported in brackets. Controls include child age and gender. $^{*}p < 0.05$

Table 3.10: Gender and Family Size

Outcome variable: total number of children	(1)
First born is girl	0.269***
	(0.010)
Odisha	-0.096***
	(0.037)
First born is girl × Odisha	-0.112**
	(0.055)
Mean	2.673
	(1.109)
Observations	37302

Notes: The outcome variable is total number of children. Controls include maternal education, household wealth, urban/rural indicator and religious group dummies. The estimation sample only includes mothers with completed fertility. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance. Source: NFHS-4.

Appendix 3.D Appendix Figures

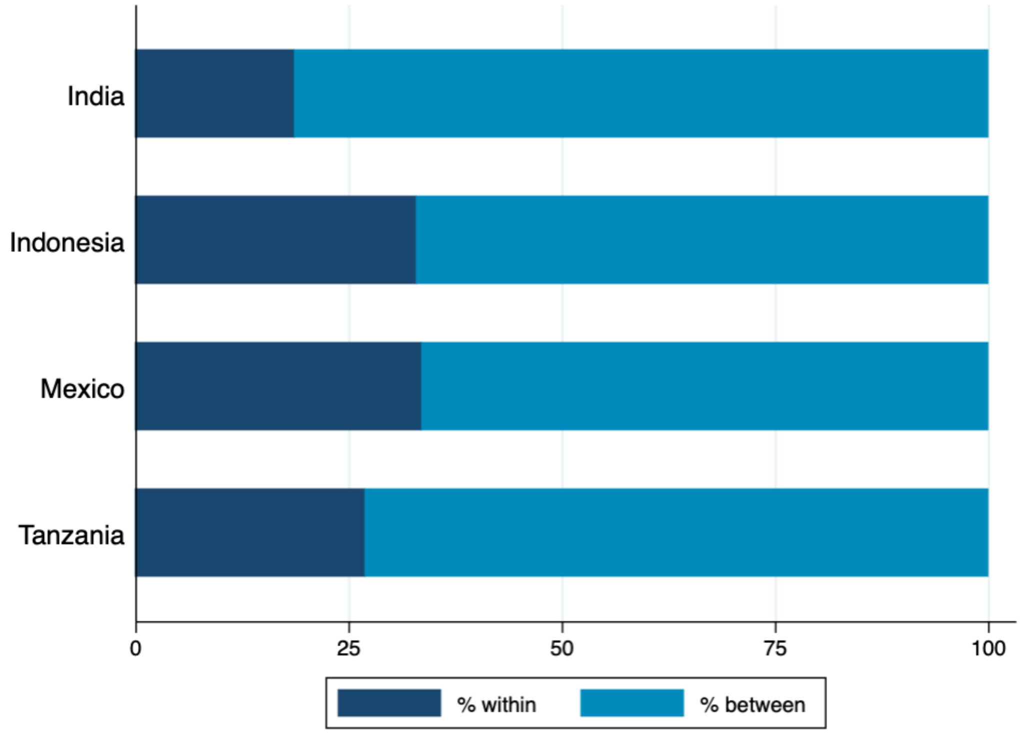


Figure 3.4: Inequality in Child Human Capital (Test Scores)

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.

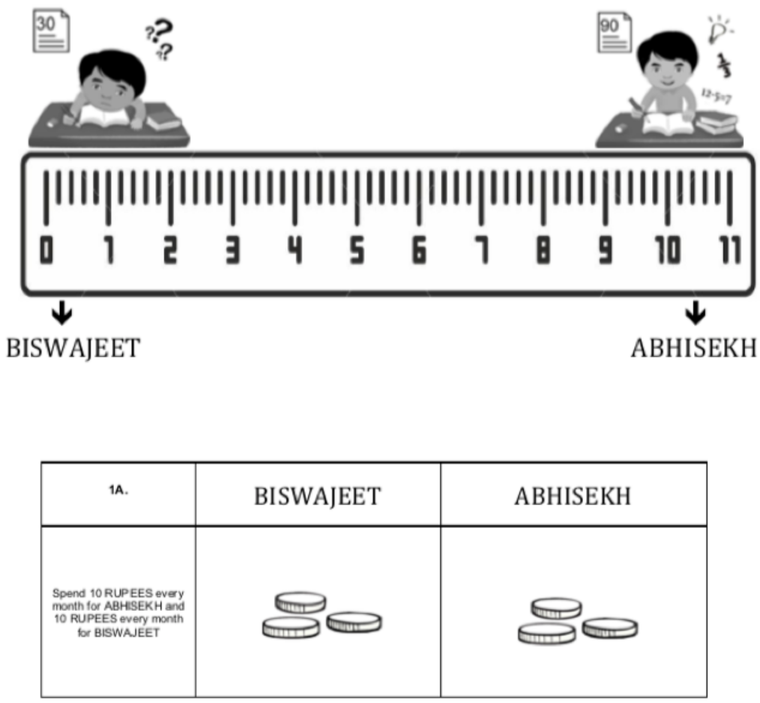
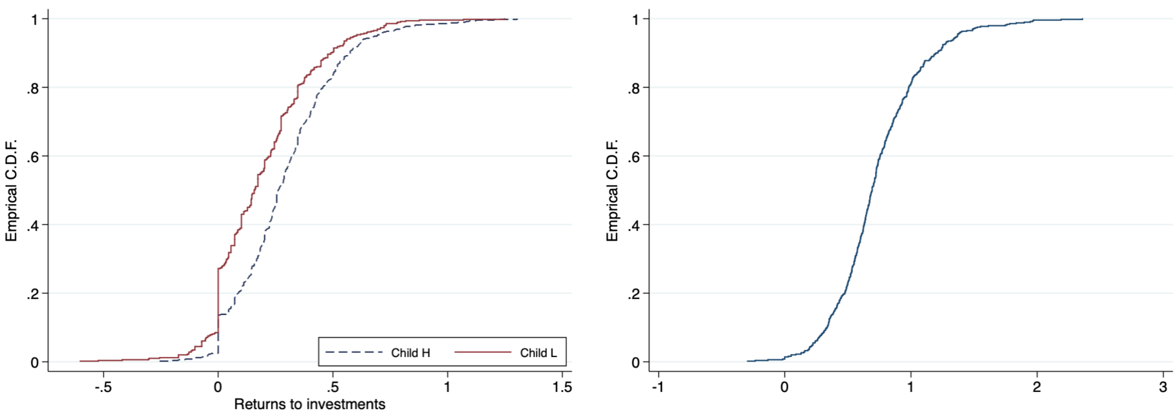


Figure 3.5: Visual Aid

Notes: This figure shows an example of visual aid used to elicit parental beliefs about the human capital production function. Child schooling ability (corresponding to endowments in the theoretical model) where 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 one child is described has having a low initial skill level, while the other child as having a high initial skill level. The level of investments in each child is low.



(A) Returns to investments

(B) Returns to endowments

Figure 3.6: Heterogeneity in Perceived Returns

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

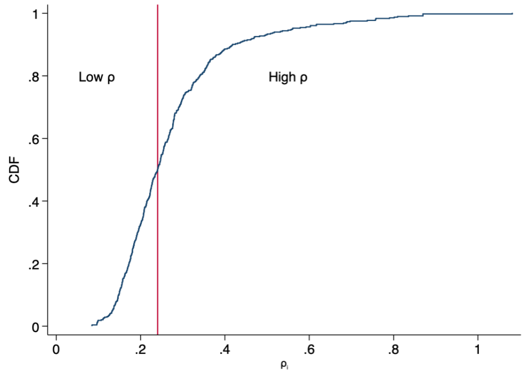


Figure 3.7: Heterogeneity in Preferences

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.

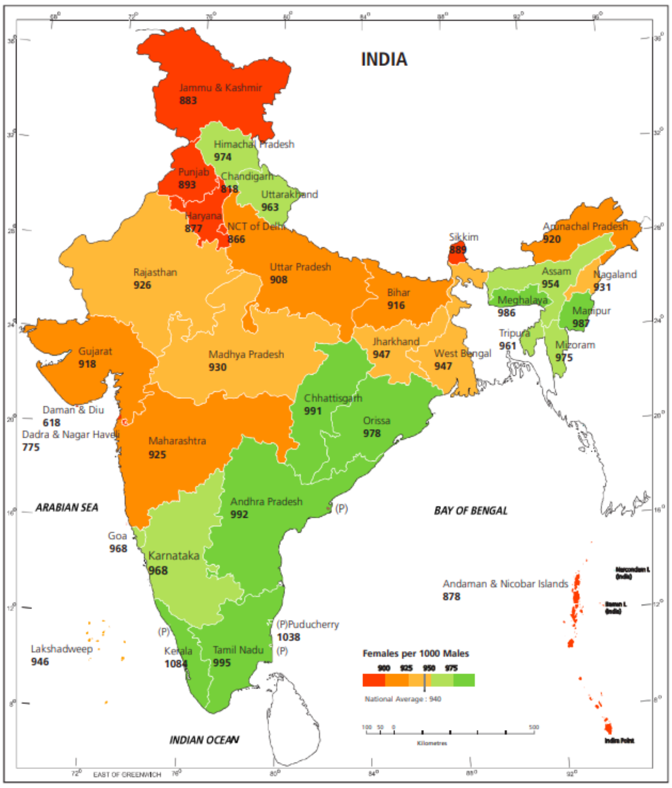
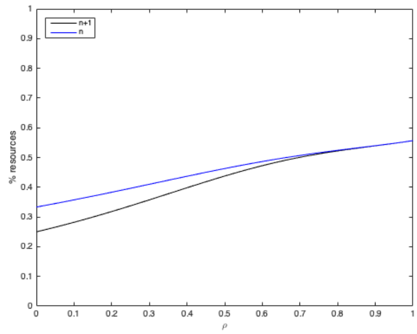
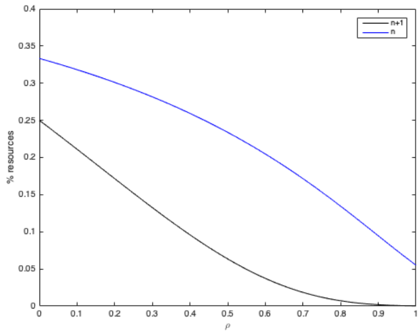


Figure 3.8: Sex Ratio in India

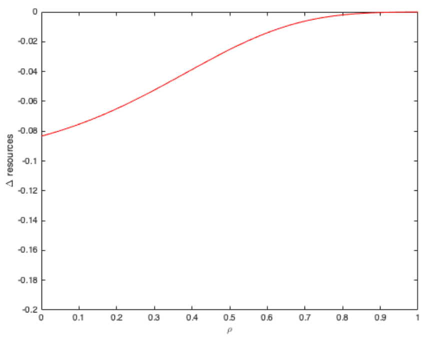
Notes: This figure plots the sex ratio (girls per 1000 boys) across Indian states. Green areas represent states where the sex ratio is higher, while red and orange areas represent states with a lower sex ratio. The figure was downloaded from https://censusindia.gov.in/2011-prov-results/data_files/india/Final_PPT_2011_chapter5.pdf on the 10/09/2020. Source: Indian Census, 2011.



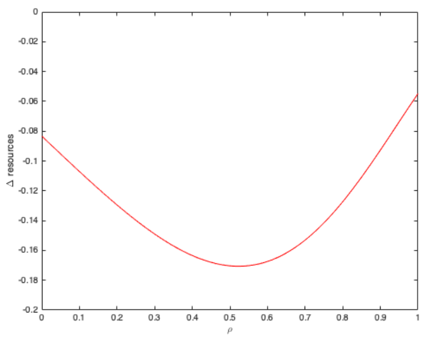
(A) Share of total resources - θ_H



(B) Share of total resources - θ_L



(C) Change in shares - θ_H



(D) Change in shares - θ_L

Figure 3.9: Family Size, Preferences and Investments

Notes: The top two panels plot the share of total resources devote to child θ_H (panel A) and θ_L (panel B) as a function of ρ for families with n children (in blue) vs $n+1$ children (in black). On the x-axis there is ρ, while on the y-axis there is the share of total household resources. The middle two panels plot the corresponding change in shares as we move from a family with n children to a family with $n+1$ children as a function of ρ, separately for child θ_H (panel C) and child θ_L (panel D). On the x-axis there is ρ, while on the y-axis there is the change in shares.

Appendix 3.E Robustness Checks for Figure 3.2

This section provides several robustness checks for the relation between fertility and the distribution of human capital in the family shown in Figure 3.2.

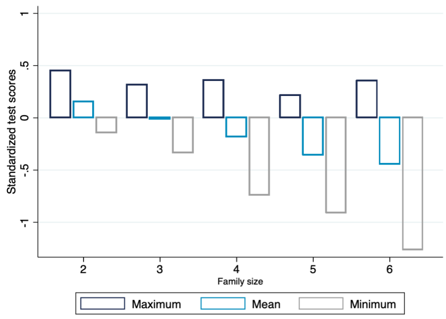
- Figure 3.10 shows the relation between family size and the distribution of child quality using age standardized test scores as measure of quality. Each sub-plot represents a different country. The figure shows that the relation in Figure 3.2 holds across countries.

- Figure 3.11 shows the relation between family size and the distribution of child quality using years of schooling as measure of quality. Each sub-plot represents a different country. The figure shows that the relation in Figure 3.2 holds across countries and is robust to the definition of child quality used.

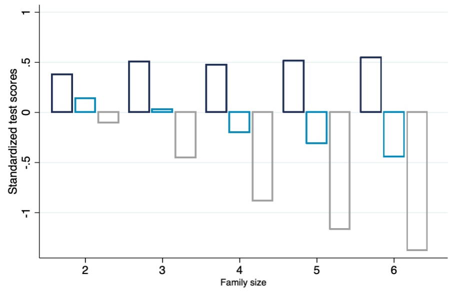
- Table 3.11 report the regression results using age standardized test scores as measure of quality. In the table, I report the results of separate regression for the mean (columns 1 to 4), the maximum (columns 5 to 8) and the minimum (column 9 to 12). Columns 1, 5 and 9 include a linear indicator for 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 3.12 report similar regression results as in Table 3.11, 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 3.11: there is a strong negative gradient in the minimum and a shallow gradient in the maximum.

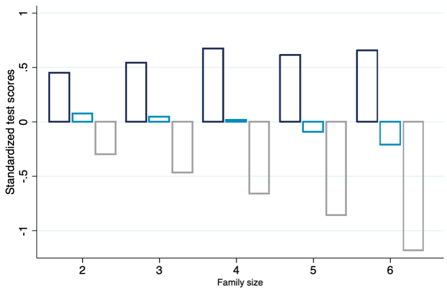
- Table 3.13 reports the IV results using years of schooling as measure of quality. Family size is instrumented using twin birth as an instrument for total family size. In the table, I report separate regressions for the mean, the maximum and the minimum. Panel A reports the results for India, while panel B reports the results for the other developing countries shown in Figure 3.11. 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 3.11. 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.



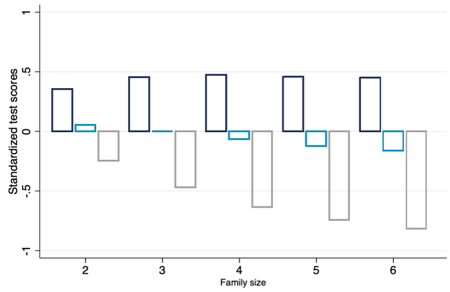
(A) India



(B) Mexico



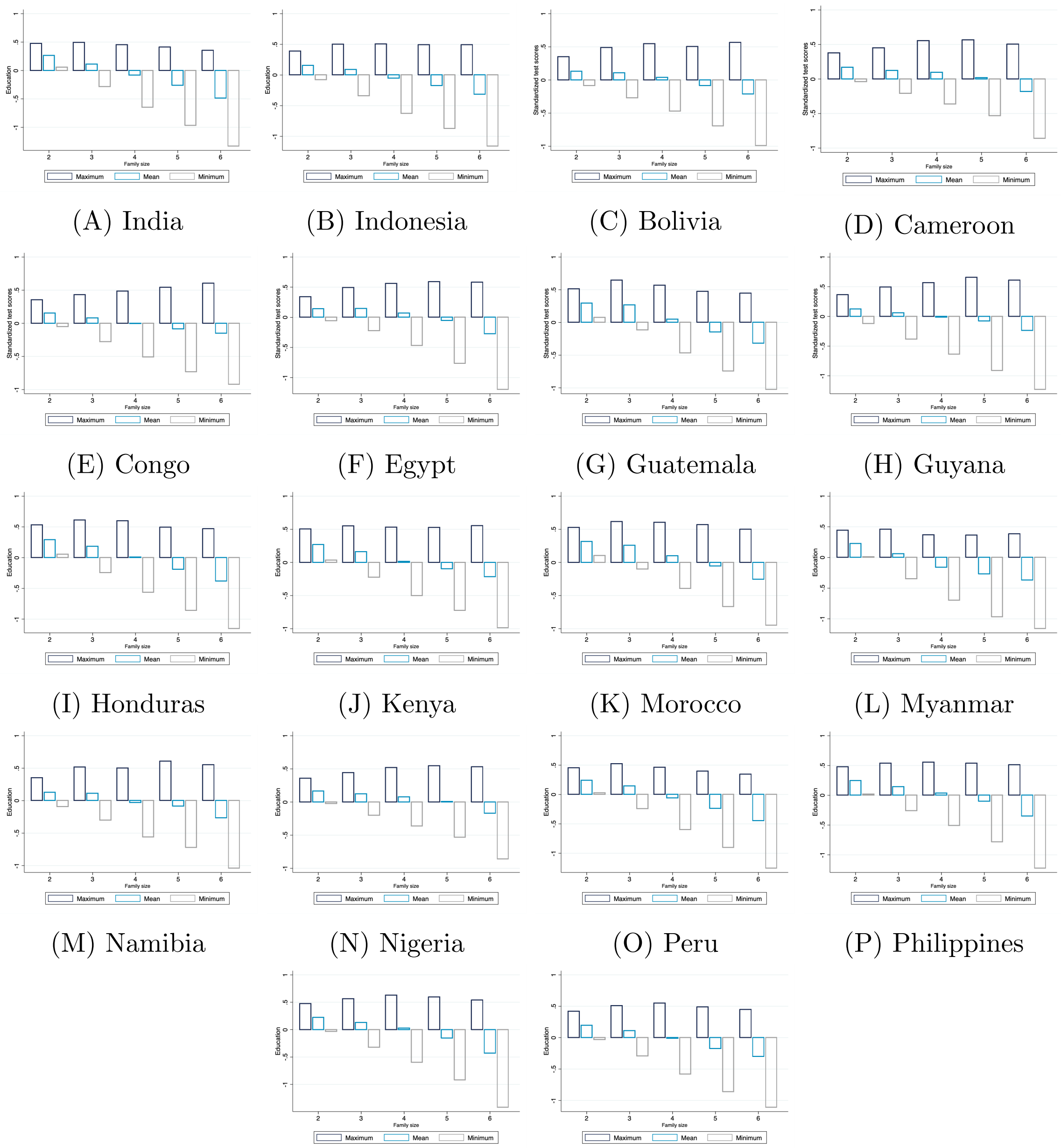
(C) Indonesia



(D) Tanzania

Figure 3.10: 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.



(A) India (B) Indonesia (C) Bolivia (D) Cameroon

(E) Congo (F) Egypt (G) Guatemala (H) Guyana

(I) Honduras (J) Kenya (K) Morocco (L) Myanmar

(M) Namibia (N) Nigeria (O) Peru (P) Philippines

(Q) Turkey (R) Zimbabwe

Figure 3.11: Fertility and Inequality in Child Human Capital (Educational Attainment)

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 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 Survey (DHS).

Table 3.11: Effect of Fertility on the Distribution of Human Capital in the Family

	Mean				Maximum				Minimum			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: India												
Linear family size	-0.149*				-0.063*				-0.266*			
	(0.011)				(0.015)				(0.017)			
Family size dummies												
3 children		-0.146*	-0.110*	-0.002		-0.134*	-0.150*	-0.053		-0.192*	-0.110*	0.023
		(0.025)	(0.027)	(0.026)		(0.030)	(0.033)	(0.032)		(0.033)	(0.036)	(0.035)
4 children		-0.296*	-0.241*	-0.019		-0.091	-0.100	0.109		-0.597*	-0.495*	-0.239*
		(0.039)	(0.043)	(0.042)		(0.055)	(0.061)	(0.060)		(0.060)	(0.068)	(0.065)
5 children		-0.459*	-0.384*	-0.142*		-0.235*	-0.253*	-0.018		-0.767*	-0.620*	-0.356*
		(0.042)	(0.050)	(0.049)		(0.060)	(0.071)	(0.070)		(0.066)	(0.081)	(0.077)
6 children or more		-0.577*	-0.431*	-0.166*		-0.100	-0.170	0.082		-1.121*	-0.821*	-0.508*
		(0.067)	(0.087)	(0.084)		(0.112)	(0.140)	(0.136)		(0.123)	(0.167)	(0.157)
F-test		48.66	19.19	2.87		6.83	6.57	2.51		65.46	23.66	10.18
p-value†		0.00	0.00	0.02		0.00	0.00	0.04		0.00	0.00	0.00
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 5% significance.

Table 3.12: Effect of Fertility on the Distribution of Human Capital in the Family - Completed Fertility Sample

	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 5% significance.

Table 3.13: Effect of Fertility on the Distribution of Human Capital in the Family - IV

	OLS			IV		
Panel A: India	Mean	Maximum	Minimum	Mean	Maximum	Minimum
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 3.11 together. * denotes 5% significance.

Appendix 3.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.

During primary school, the parents decide how much money to spend on educational resources that will help Abhisekh and Biswajeet acquire new skills and progress in their education (e.g. books, private tuitions etc.). Remember that Abhisekh is among the top three students in his school and Biswajeet is among the bottom three students in his school.

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.

Chapter 4

Helping Struggling Students and Benefiting All: Peer Effects in Primary Education

4.1 Introduction

Despite the dramatic increase in primary enrollment rates in the last 50 years, it is estimated that over 125 million children worldwide are not acquiring functional literacy skills, even after four years in school (World Bank (2018)). Adequately serving the learning needs of the lowest achieving children poses a challenge to modern education systems (United Nations (2016)). This is in part because schools group children by age rather than learning levels, which often forces teachers in large classrooms to address heterogeneous instructional needs (Muralidharan, Singh, & Ganimian (2019)). Not surprisingly, there is a long-standing concern among parents and educators that low-achieving students can negatively affect their classmates' learning either directly, or indirectly by diverting resources away from the rest of the class (Duflo, Dupas, & Kremer (2011), Reback (2008)). This paper quantifies the effect of providing support to low-achieving students on the academic performance of their higher-achieving peers.

In the more than 50 years since the publication of the Coleman Report (Coleman (1968)), a large body of research in economics, education and sociology has documented the central role played by peers in determining academic outcomes at all education levels.[1] Of particular interest is the effect that disruptive students can have on the rest of their classmates. Theoretically, the *bad apple model* emphasizes how the presence of even a single disruptive student can slow down learning for the rest of

[1]The Coleman Report concluded that much of the achievement gap between white and black students could be attributed to differences in the composition of peers these students faced in American public schools. For studies that analyze peer effects in elementary and secondary schools, see, for example, Hoxby (2000), Hanushek, Kain, Markman, & Rivkin (2003), and Whitmore (2005). See Sacerdote (2001), Zimmerman (2003) and Stinebrickner & Stinebrickner (2006) for evidence at the university and college levels.

the class (Lazear (2001)). The concept of *bad apple* has been operationalized in the empirical literature using different measures: repeaters (Lavy, Paserman, & Schlosser (2011)), children exposed to domestic violence (Carrell & Hoekstra (2010)), boys with names most commonly given to girls (Figlio (2007)). These papers document a large and negative effect of *bad apples* on their peers' academic performance. Moreover, there is evidence that these effects are persistent and translate into lower educational attainment and reduced earnings (Carrell, Hoekstra, & Kuka (2018)).

Equity considerations aside, the existence of negative externalities caused by low-achieving students should in itself provide a compelling justification that underscores why *all* parents and policy makers should be concerned about how to properly support this group of students. Yet, the previous literature has limited itself to describing the phenomenon rather than studying potential policies that could attenuate the impact of low-achieving peers. Designing policies that could effectively tackle this issue requires a deeper understanding of the mechanisms underlying this negative externality. For instance, a *bad apple* may take away instructional time by asking questions to which all other students already know the answer. Alternatively, she may cause classroom disruption through misbehaviour, forcing the teacher to allocate time away from learning activities. Evidence shows that low-achieving students are often the very same students who cause classroom disruption (Lavy, Paserman, & Schlosser (2011)), so that disentangling these two mechanisms becomes problematic. Nonetheless, this distinction is crucial in order to determine whether interventions that improve the academic performance of low-achieving students without targeting miss-behaviour can improve learning for the rest of the class.

In this paper, we exploit a randomized evaluation of a remedying education program that targeted struggling students, to study whether an exogenous improvement in the test scores at the bottom of the class can generate gains at the top. The intervention we consider aimed to improve reading among low-achieving third-grade students in Colombia. At the beginning of the school year, all students were tested to determine their baseline literacy level. Students with baseline reading scores lower than a certain threshold were deemed eligible to receive the tutoring classes.[2] Schools were then randomized into treatment and control groups. Eligible children in treatment schools were taken out of the regular classes to work in small groups with a qualified tutor, who followed a structured pedagogical curriculum for 40 minutes, three times a week. Eligible children in control schools continued their classes as usual. The intervention improved literacy skills of low-achieving students by one third of a standard deviation (Alvarez Marinelli, Berlinski, & Busso (2019)).

Importantly, the research design naturally generates two groups of students within the same class: low-achieving students who were *eligible* to receive the intervention, and higher-achieving students who were *not eligible* (henceforth, we refer to the eligible

[2] This eligibility threshold was determined by local pedagogues. It was based on the skill level expected from a second-grade pupil.

students as low achievers, and to their classmates as high achievers). Determination of students' eligibility for the tutoring program took place prior to schools' randomization into treatment and control status, allowing us to identify these two groups of children both in treated and control schools.

We find that *non-eligible* children in treated schools scored 0.108 of a standard deviation higher than similar children in the control group. This coefficient is sizable and represents roughly 30 percent of the treatment effect we measure on the *eligible* students. This result is economically meaningful, and its magnitude can be compared to a more commonly proposed school-level reform, tracking by prior achievement (Duflo, Dupas, & Kremer (2011)).

We interpret the reduced-form effect on high-achieving students as a spillover effect within treatment units, and estimate linear-in-means models of peer effects. Credibly identifying peer effects is challenging given the well-known issues of selection, reflection, and correlated unobservables (Manski (1993)). We overcome these identification challenges by exploiting the experimentally induced variation in the outcome of a sub-set of individuals in the peer group. This approach is defined by Moffitt (2001) as a *partial population experiment*. Randomization of the program solves the reflection problem as it induces exogenous variation in the outcomes of low-performing children without *directly* affecting high-ability students. Second, random assignment implies that the treatment is orthogonal to all observables and unobservable characteristics, solving the problem of correlated unobservables. Finally, as peer groups are established before the policy change and remain fixed throughout the experiment, endogenous group membership is not an issue either. We can think of peer effects as being conditional on any selection into groups that might have taken place prior to the experiment.

We find strong evidence of peer effects in academic outcomes. A one-standard-deviation increase in peers' *contemporaneous* test scores increases individual reading score by 0.679 of a standard deviation. We also find evidence of non-linearities, with significantly stronger effects for students at the top of the achievement distribution compared to students at the bottom.

The term *peer effect* is generally used as an umbrella term that comprises any externality implying that peers' outcomes affect an individual's outcome. Potential mechanisms effectively include peer-to-peer learning, student misbehaviour, and teacher practices (Sacerdote (2011)). In this paper, however, we seek to distinguish between these alternative mechanisms because such distinctions might be key for the design of optimal education policies. Using survey data on teachers, we show that teachers did not change their practices in response to the program. At the same time, we find suggestive evidence that a reduction in classroom disruption may drive some results. This suggests that low levels of achievement foster disruptive behaviour, and that interventions that only target learning without modifying behaviour can relax the constraints posed by low-achieving students on the rest of their classmates. We further rule out an alternative explanation coming from a reduction in class size, to which *non-eligible* children were only marginally exposed.

This paper stands out from the literature on peer effects in education in a number of fundamental ways. First, in contrast with most of the previous literature, we study the impact of peers' contemporaneous achievement – the *endogenous* effect in the terminology of Manski (1993) – on individual outcomes directly, as opposed to peers' background characteristics, such as gender, race, or prior achievement.[3] This is particularly important given that research demonstrates that once one controls for peers' achievement, these background characteristics do not matter for student outcomes (see Hoxby & Weingarth (2005)). Moreover, peer effects stemming from background characteristics do not entail a social-multiplier effect (Sacerdote (2011)). On the other hand, effects stemming from peers' contemporaneous achievement have the potential to generate social-multiplier effects. In our setting, the beneficial effects of improving the academic achievement of low-achieving students spill over onto non-treated students, magnifying the total output of the program.

Second, we focus on peer effects in naturally occurring groups, and exploit the random variation in the outcomes of a subset of group members. This distances our work from that strand of the literature that uses the random allocation of students to groups. This distinction is particularly important given that opportunities to randomly assign peers are rare in real-world settings, whereas the possibility of randomly treating a subset of individuals within a group might not be so rare.[4] Moreover, a particularly important issue is whether the results in those studies that exploit the random allocation of peers are generalizable to naturally occurring peer groups. The results in Carrell, Sacerdote, & West (2013) directly speak to this issue by highlighting how exogenously manipulating group composition might have unpredictable (and sometimes detrimental) effects on students' academic outcomes. In the context of the U.S. Air Force Academy, Carrell, Sacerdote, & West (2013) show that low-ability students placed into "optimally" designed peer groups perform significantly worse than comparable students who were randomly allocated to squadrons.[5] The explanation for this result is that the treatment changed the patterns of social interactions in ways that were key for student achievement. This evidence highlights how policy-induced patterns of social interactions may be a major obstacle to predicting the effects of altering peers' composition. Such concerns cast some doubt on the external validity of studies that randomly assign individuals to groups.

---

[3]An important exception is Fruehwirth (2013), who estimates spillover effects in academic outcomes in the context of a student accountability policy in North Carolina.

[4]See Sacerdote (2001), Cullen, Jacob, & Levitt (2006), Lyle (2007), Carrell, Fullerton, & West (2009), Duflo, Dupas, & Kremer (2011), and Carrell, Sacerdote, & West (2013) for examples of papers that use the random allocation of students to groups to estimate peer effects.

[5]In the U.S. Air Force Academy, incoming students are randomly allocated to squadrons. The design of "optimal" peer groups relied on estimating flexible reduced-form specifications of peer-effects in academic achievement using pre-treatment data. The objective was to maximize the outcomes of low-performing students.

Third, this study provides the first successful example of how peer effects can be exploited in the design of public policies aimed at improving students' academic performance. Differently from Carrell, Sacerdote, & West (2013), who focus on *exogenous* peer effects by randomly varying the *composition* of peer groups, we exploit the existence of *endogenous* effects within pre-existing peer groups. Our results show that policies aimed at improving the bottom of the distribution have the potential to generate social-multiplier effects. Importantly, the findings indicate that it is possible to substantially improve academic outcomes *for all* with interventions targeted to the weakest. We believe that these considerations are important to inform any policy debate concerned with the allocation of public funds to education.

Finally, by showing how the failure to consider general equilibrium effects might lead to an *underestimation* of the impacts of a policy, this paper also contributes to the policy evaluation literature. It is important to underscore that in our context, confining the consideration of the treatment effect to the *eligible* population would underestimate the benefits of the program by 47 percent. Thus, our findings underline the need to collect data on the entire local economy to fully appreciate policy effects. In addition, the results suggest the importance of experimentally manipulating individuals' treatment status within treatment units (schools in our setting) to identify social interactions.

The rest of this paper is organized as follows: Section 4.2 describes the remedying education intervention, the evaluation design as well as the experimental results on the sample of low achievers. In Section 4.3, we discuss the issues related to the identification of peer effects and explain how we use the intervention to overcome these identification challenges. Section 4.4 presents the result. Section 4.5 addresses potential threats to identification, and discusses mechanisms and policy implications. Section 4.6 concludes.

4.2 The Remedial Literacy Program

4.2.1 Setting

The remedial education program took place among third-grade students in public elementary schools in the municipality of Manizales in Colombia during three consecutive years (2015-2017).[6] The municipality scored slightly above the national mean among third-graders in the 2016 national standardized language achievement tests (Pruebas

[6] Manizales is a midsize city in central Colombia. Approximately 13.8 percent of its residents have incomes below the poverty line, compared to a national average of 27 percent (Departamento Administrativo Nacional de Estadstica (2018)). About 6.9 percent of the municipality's residents lived in rural areas. More than 18,000 children were enrolled in the first five grades of the public elementary school system. In our sample, 97 percent of students fall in the first three levels of the social stratification classification scale used to target social programs in Colombia. In Manizales, about 78 percent of school-aged children attend public schools, and most children in our sample attended the school closest to their home.

Saber). However, almost 45 percent of students scored at or below the minimal-knowledge threshold in standardized official tests (Alcadía de Manizales (2017)). As a result, the local Secretary of Education, in partnership with a local NGO (Fundacion Luker) and the Inter-American Development Bank, implemented a remedial program to improve reading fluency among struggling third-grade students, and designed the evaluation of its effectiveness.[7]

4.2.2 Small-group Tutorials for Low-achieving Students

The program provided students with 40-minute sessions three times a week for up to 16 weeks in the second half of the school year. The tutorials were conducted in small groups (six students maximum) and followed a simple structure. During each lesson tutors explained the objectives and activities, modeled the different exercises, and used both guided practice and student independent practice. The sessions used a curriculum designed and refined by international experts with support from a local team. The curriculum was based on a phonics approach. Lessons emphasized the ability to identify and manipulate units of oral language, the ability to recognize letter symbols and the sounds they represent, the ability to use combinations of letters that represent speech sounds, reading of words, and reading fluency of sentences and paragraphs. It also worked on vocabulary and strategies for reading comprehension.

The intervention targeted struggling readers who were identified using a measure of language development. At the beginning of the school year, the Early Grade Reading Assessment (EGRA) was used to collect information on the following literacy subtasks: knowledge of letter sounds, reading of non-words, fluency of oral reading, and reading comprehension.[8] The tests were applied to the universe of children in public schools. This information was used to determine which students were eligible to participate in the tutorials in each school.[9] Importantly for this paper, this strategy naturally generated two groups of students within the same class: a group of low-achieving students who were *eligible* to receive the intervention, and a group of higher-achieving students who were *not eligible*.

Throughout the paper, we define two students as peers if they attend the same class in the same school, rather than by grade level, as many existing studies have done. We argue that the classroom-based peer definition offers a better approximation of how students interact in primary schools. For example, children in our sample spend at least 6 hours a day for roughly 165 days a year with their classmates, while the

[7] Alvarez Marinelli, Berlinski, & Busso (2019) provide further details regarding the intervention, the experiment, and its results on the target population.

[8] For more information on the EGRA test see Dubeck & Gove (2015).

[9] The eligibility criteria were different between the first and the subsequent cohorts. For the first cohort, students were eligible if they scored in the bottom 25 percent of an equally weighted composite index of the following EGRA subtasks: reading of non-words, fluency of oral reading, and reading comprehension. The eligibility criteria were changed during the second and third cohorts, when eligibility required that children correctly read fewer than 60 of the 132 words in a paragraph in the EGRA fluency of oral reading subtask.

occasions to interact with other schoolmates are rather rare and mostly limited to playtime during recess. This is particularly important given that peer-effects estimates have been shown to greatly depend on the accuracy of the identification of relevant peers (Carrell, Fullerton, & West (2009)). For instance, Burke & Sass (2013) find evidence of peer effects at the classroom level but not at the grade-within-school level for elementary school children.

After collecting baseline data on children, schools were randomly assigned to treatment in the following ways: i) Schools were sorted based on how many low-achieving students were enrolled in third grade, and stratified in blocks of size two. ii) Within these strata one school was randomly selected to receive treatment, and the other one to be a control. Low-achieving students in treatment schools participated in the remedial reading program, while those in the control schools carried on with their usual classroom learning experiences. iii) Tutors were then randomly assigned to schools (one tutor per school), and, in schools with more than six eligible children, students were assigned to equally sized tutorials.[10] This procedure was repeated each year of the intervention so that the *same* school could potentially be in a different treatment status from one year to the next.

Struggling readers in treatment schools were taken out of the classrooms during regular school hours. Tutorials took place in a designated school space at different times during the school day, not specifically at the same time as literacy classes. Tutors were hired each year of the intervention, and trained specifically to deliver the remedying program. They were trained primary school teachers, psychologists, or audiologists with some teaching experience. There were no planned interactions between tutors and regular classroom teachers or non-eligible students.

The sessions took place for 120 minutes each week. During that time the low-achieving students participated in the remedial intervention, and their high-achieving peers continued receiving instruction using the standard curriculum. The high-achieving students had no direct interaction with the remedial intervention program. Their only exposure was indirect, occurring through their interactions with low-achieving children with whom they shared the classroom every day and in all subjects.

4.2.3 Data and Descriptive Statistics

The key outcome of interest is student achievement as measured by standardized math and language test scores on the Early Grade Reading Assessment and the Early Grade Math Assessment (EGRA and EGMA, RTI-International (2009)). Both tests were administered at the end of the school year by trained enumerators, who interviewed all students individually using a tablet. The application of the tests took less than 20 minutes per student. Our main outcome variable is the sum of correct answers across

[10] In the case of the first cohort, when there were more than six low-achieving students in the school, schools organized the compositions of the tutorials. In the second and third cohorts, eligible students were randomly assigned to tutorials.

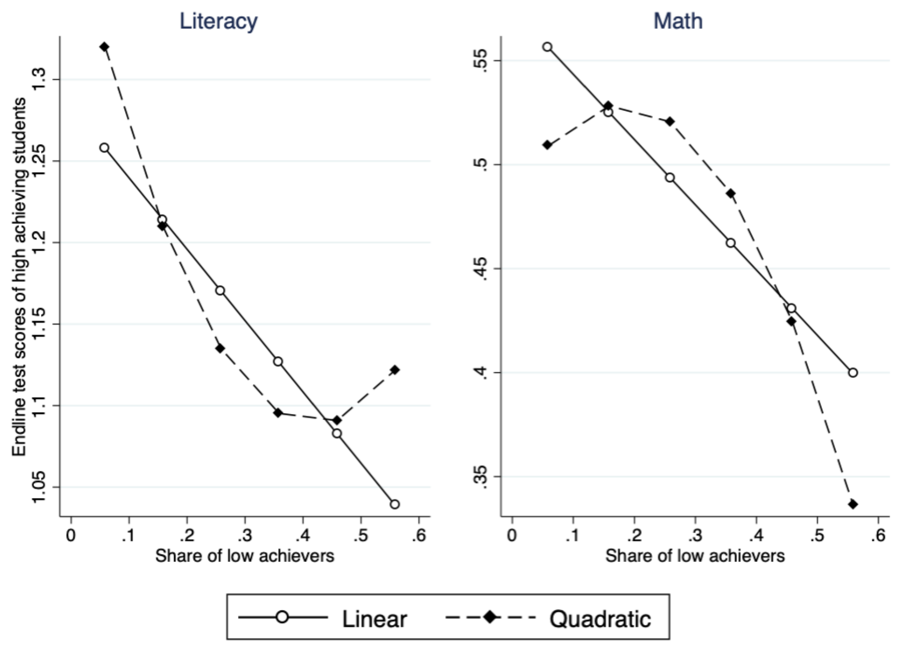


Figure 4.1: Linear and Quadratic Fits of End-line Scores of High-achieving Students by Classroom Share of Low Achievers (Control Schools Only)

Lines represent linear and quadratic fits of standardized end-line test scores of high-achieving students as a function of baseline share of low achievers in non-treatment schools. Controls include a second-order polynomial in age, gender and school fixed effects. The figure is trimmed at the 5th and 95th percentiles of the distribution of classroom share of low achievers.

all reading and math subtasks standardized by the mean and standard deviation observed in the control group of each cohort. We also report similarly defined literacy and math scores. In addition, we rely on information on child gender, age, and socio-economic status extracted from the administrative school records of the Integrated Enrollment System (Sistema Integrado de Matricula, SIMAT), the national database for the registration of students in public education in Colombia. This dataset also contains school-level information, which we used to compute class size.

Consistent with the evidence in Carrell & Hoekstra (2010) and Lavy, Paserman, & Schlosser (2011), in our sample we find that the test scores of high-achieving students negatively correlate with the share of low achievers in their classroom. Figure 4.1 plots the relationship between end-of-the-year literacy and math scores of non-eligible students and share of low achievers prior to the intervention, in the sample of control schools only. Both for literacy and math, average achievement decreases monotonically with the share of low-achieving students.[11]

Because the goal of this study is to understand the role that peers play in standard classroom settings (medium-sized, single-grade classrooms with one teacher), we restrict the analysis to randomization strata in which enrolment in third grade in

[11] While there might be unobservable characteristics that simultaneously affect both the share of low achievers in the class and the test scores of their higher-achieving peers, the figure controls for that type of selection by including school fixed effects.

Table 4.1: Baseline School and Individual Characteristics by Treatment Group

	Treatment		Control		p-value Treatment = Control
	Means	S.D.	Mean	S.D.	
Panel A	*School and class characterisics*				
Rural	0.026	0.160	0.000	0.000	0.326
Class size	30.783	6.891	30.780	6.055	0.969
Eligible share	0.254	0.145	0.280	0.154	0.428
Panel B	*Individual characteristics - Low achieving students*				
Age	8.563	0.943	8.569	0.987	0.922
Gender (girl)	0.490	0.500	0.515	0.500	0.437
Disability (no)	0.823	0.382	0.819	0.385	0.842
SES	0.298	0.458	0.319	0.466	0.488
Literacy score	93.068	23.625	92.517	26.328	0.773
Math score	20.981	7.694	20.306	8.048	0.195
Total score	114.049	27.119	112.823	29.724	0.587
Panel C	*Individual characteristics - High achieving students*				
Age	8.435	0.874	8.381	0.820	0.155
Gender (girl)	0.515	0.500	0.505	0.500	0.742
Disability (no)	0.898	0.302	0.906	0.292	0.460
SES	0.233	0.423	0.219	0.413	0.497
Literacy score	139.376	28.057	140.447	28.371	0.616
Math score	25.757	8.072	25.487	7.782	0.508
Total score	165.133	30.751	165.935	30.929	0.713

Notes: Panel A: school characteristics. Panels B and C: individual characteristics. p-values are for tests of equality of the means across treatment and control groups.

both schools in the stratum was larger than 20 students. Table 4.1 presents summary statistics of the schools and children in our sample by treatment status. Panel A reports school and class characteristics. Children's characteristics are reported separately for low-achieving students, who were eligible to take part to the intervention, and high-achieving students who were not eligible, in panels B and C respectively.

The treated and control groups are very similar in terms of observable characteristics, as one would expect given the randomized nature of the program. Low-achieving students and higher-achieving students do not seem to differ much in terms of observable characteristics other than one: their school achievement levels. The scores of eligible children are significantly lower than those of non-eligible children. The magnitude of these differences is 47 points in literacy (out of 240, with p-value = 0.000) and 5 points in math (out of 52, with p-value = 0.000). For literacy, this knowledge gap is comparable to a full year of learning in the control group

Table 4.2: Treatment Effects - Low Achieving Students

	(1)	(2)	(3)	(4)
Panel A: Literacy	0.362	0.361	0.358	0.342
	(0.091)	(0.091)	(0.091)	(0.078)
Panel B: Math	0.092	0.081	0.079	0.131
	(0.060)	(0.059)	(0.061)	(0.057)
Panel C: Total score	0.317	0.314	0.312	0.315
	(0.083)	(0.083)	(0.083)	(0.072)
Observations	1889	1889	1889	2413
Individual controls		✓	✓	✓
Class controls			✓	

Notes: The outcome variables are standardized test scores. Individual controls include a second-order polynomial in age and gender. Class control include class size and number of classrooms in the schools. All regressions control for school fixed effects. Robust standard errors are clustered at the classroom level, and presented in parentheses. Column 4 reports the results in the evaluation sample in Alvarez Marinelli, Berlinski, & Busso (2019).

4.2.4 Experimental Results

Alvarez Marinelli, Berlinski, & Busso (2019) present the evaluation of the program among the population of eligible students, showing that, immediately after the experiment finished, the overall literacy score of low-achieving students in treated schools improved compared to control students. Table 4.2 replicates the main experimental results. At end line, the scores of low-achieving students in treatment schools were 0.362 of a standard deviation higher in literacy than those of similar, low-achieving students in control schools (column 1, Panel A of Table 4.2). The coefficient is virtually unchanged when including individual or class-level control variables (columns 2 and 3). Column 4 shows the results in the complete (unrestricted) sample. Panel B of Table 4.2 also reports a positive but not statistically significant result on math scores. The results in Table 4.2 clearly highlight that the intervention was very effective in increasing the test scores of low achievers. This is important because it provides us with a source of exogenous variation in peers' contemporaneous test scores that we will exploit to study peer effects in academic achievement (see Section 4.3).

Alvarez Marinelli, Berlinski, & Busso (2019) further show that the effect of the intervention is homogeneous in key respects. There are fairly constant quantile treatment effects. There seems to be no significant heterogeneity among students who attended smaller or larger tutorials, or those who had comparatively worse or better tutorial peers, or those who were in more homogeneous or more heterogeneous tutorial groups, in terms of baseline reading ability.

4.3 Identification Strategy and Methodology

As discussed in Manski (1993), credibly identifying and quantifying peer effects poses important empirical challenges. In this section, we first describe these challenges, and then explain in detail our identification strategy.

First, the *simultaneity or reflection problem* arises as students affect each other, so that there is no exclusion restriction that can be used to distinguish the effect that the individual has on the group from the effect of the group on the individual herself. Second, correlated unobservables plague identification when not all relevant group or individual characteristics are observed. These unobservables can generate a spurious correlation in outcomes that do not represent causal effects (Lyle (2007)).[12] Third, endogenous group membership is an issue because individuals self-select into peer groups or classrooms in a manner that is unobserved by the researcher. Positive selection frequently occurs with similar people joining the same group. This phenomenon, known as *homophily,* implies an upward bias in the estimated magnitude of peer effects.[13]

Previous research has tried to overcome these issues by including an extended set of controls for students and school characteristics. This often means using student and school fixed effects, or exploiting the naturally occurring variation in cohort composition over time within a school to deal with selection into peer groups.[14] Because results could still be biased, a second set of studies has exploited the random or quasi-random variation in peer group composition to identify peer effects.[15] While these papers credibly tackle the issue of self-selection, they effectively answer the question, "What would happen if individuals were randomly assigned to peer groups?". Whether the findings of these studies are generalizable to naturally occurring peer groups is not obvious. In particular, the patterns of social interactions that exist in these two different types of groups may differ fundamentally, resulting in different effects of peers on individual outcomes.

This is not a mere theoretical speculation. The results in Carrell, Sacerdote, &

[12]This would still be a problem in individuals were randomly assigned to groups. For example, in the educational context, this could be interpreted as a teacher fixed effect. Randomization of students into classes would still imply that students within the same class are exposed to the same teacher: a positive correlation in outcomes could be the results of same teacher exposure rather than a causal effect of peers.

[13]McPherson, Smith-Lovin, & Cook (2001) report the relevance of this phenomenon to explain the formation of social ties in a wide range of contexts including marriage, work advice, information transfer, and friendship. In the educational context, Carrell, Sacerdote, & West (2013) report that students are more likely to interact with peers of similar ability and form homogeneous sub-groups within the class, even when they are randomly assigned to classes.

[14]Studies that use this strategy include Hoxby (2000), Hanushek, Kain, Markman, & Rivkin (2003), Lefgren (2004), Hoxby & Weingarth (2005), Carrell & Hoekstra (2010), Lavy, Paserman, & Schlosser (2011), Burke & Sass (2013), and Card & Giuliano (2016).

[15]Papers that use the random assignment of students to groups include Sacerdote (2001), Cullen, Jacob, & Levitt (2006), Lyle (2007), Carrell, Fullerton, & West (2009), Duflo, Dupas, & Kremer (2011), and Carrell, Sacerdote, & West (2013). Other studies have used natural experiments as source of exogenous variations in peer composition. For examples, see Angrist & Lang (2004), Cipollone & Rosolia (2007) and Imberman, Kugler, & Sacerdote (2012).

West (2013) directly speak to this issue. Using the random allocation of cadets to squadrons in the U.S. Air Force Academy, Carrell, Sacerdote, & West (2013) estimate flexible reduced-form specifications of peer effects in academic achievement. Using these estimates, they allocate incoming students to squadrons in order to maximize the achievement of lowest-performing students. Surprisingly, their findings show that low-ability students placed into these "optimally" designed peer groups performed significantly worse than comparable low-ability students who were randomly allocated to squadrons. The explanation for this puzzling result is that the treatment changed the endogenous patterns of social interactions in ways that were key for student achievement. In particular, the authors show that within their optimally designed groups, low-performing students avoided interacting with high achievers (the very students they intended them to interact with), and instead formed more homogeneous subgroups. This evidence highlights how policy-induced patterns of social interactions may be a major obstacle to predicting the effects of altering peer groups. The findings cast some doubt on the external validity of studies that randomly assign individuals to groups.

Finally, a small but growing literature exploits *partial population experiments* (Moffitt (2001)), to study peer effects in naturally occurring groups. This approach uses the experimentally induced variation in the outcomes of a subset of individuals in the relevant peer group to identify peer effects for the non-treated individuals. This approach has been used to study labor market outcomes (Hesselius, Nilsson, & Johansson (2009)), financial decisions (Bursztyn, Ederer, Ferman, & Yuchtman (2014)), retirement plan decisions (Duflo & Saez (2003)), social program participation (Dahl, Løken, & Mogstad (2014)), and healthy behaviour (Centola (2010)). Only a very few papers have used this approach in the context of education, and most of them have looked at peer effects in school enrollment rather than academic achievement (Bobonis & Finan (2009), Lalive & Cattaneo (2009) and Angelucci, De Giorgi, Rangel, & Rasul (2010)).[16] Our approach is similar to these studies in that we exploit a randomized control trial designed to improve reading fluency among low-performing students to study academic achievement of their *non-treated*, high-achieving peers. This remedying education program offers a unique opportunity to analyze whether an exogenous increase in the test scores of peers within a class increases individual achievement.

The essence of our identification strategy can be more easily understood by considering the following system of equations: For simplicity imagine that the reference group (i.e., the class) is only composed of three students: A, B and C, where A and

[16] So far as we are aware, the only other paper that uses a partial population experiment to look at peer effects in achievement is Boozer & Cacciola (2001) in the context of the Tennessee Student-Teacher Achievement Ratio experiment (Project STAR). However, that study analyzes peer effects in groups that are randomly assigned. Therefore, the concerns of external validity raised above are still valid for this study.

B are high achievers and C is a low-achieving student. Then we can write:

$$\begin{aligned} y_{A,G} &= \alpha + \rho\left(\frac{y_B + y_C}{2}\right) + \delta x_A + \gamma\left(\frac{x_B + x_C}{2}\right) + \theta\omega_G + \epsilon_{A,G} \\ y_{B,G} &= \alpha + \rho\left(\frac{y_A + y_C}{2}\right) + \delta x_B + \gamma\left(\frac{x_A + x_C}{2}\right) + \theta\omega_G + \epsilon_{B,G} \\ y_{C,G} &= \alpha + \rho\left(\frac{y_A + y_B}{2}\right) + \delta x_C + \gamma\left(\frac{x_A + x_B}{2}\right) + \theta\omega_G + \tau T_G + \epsilon_{C,G} \end{aligned} \tag{4.1}$$

where $y_{i,G}$ is the academic achievement of student i in group G, x_i are individual observable characteristics, ω_G are observable group specific characteristics, and $\epsilon_{i,G}$ is an error term. Notice that treatment T_G varies randomly across groups, but there is no change for any high-ability student. In the terminology of Manski (1993), ρ is the endogenous effect emanating from peers' contemporaneous outcomes, while γ is the exogenous effect from peers' background characteristics. The focus of this paper is the endogenous peer effect.

The random assignment of the treatment overcomes the identification challenges in the following ways: First, it solves the reflection problem because the experiment induces exogenous variation in the outcomes of the low-performing child (student C) without *directly* affecting high-ability students (A and B). Second, randomization implies that the treatment is orthogonal to all observable and unobservable characteristics (x_i, ω_G, and $\epsilon_{i,G}$), solving the problem of correlated unobservables.[17] Finally, because peer groups are established before the policy change and fixed throughout the experiment, endogenous group membership is not an issue. We can think of peer effects as being conditional on any selection into groups that might have taken place prior to the experiment.

We can identify the causal effect of the program τ by regressing $y_{C,G}$ on T_G. The endogenous peer effect ρ is identified by regressing $y_{i,G}$ (for $i = A, B$) on T_G and scaling by $\hat{\tau}$. This is equivalent to an instrumental variable strategy that uses T_G as instrument for average peer achievement in the equation of high-achieving students.

Formally, we estimate the following linear-in-means model of peer effects using two-stage least squares (2SLS) on the sample of *high achievers* only (i.e., on the sample of students who were *not eligible* for the remedying intervention):

$$y_{icst} = \rho\bar{y}_{-icst} + X_{icst}\beta + \omega_s + \lambda_t + \epsilon_{icst} \tag{4.2}$$

where i is the student, c is the class, s is the school and t is the cohort. The outcome variable y_{icst} is the test score of a student (expressed in standard deviations of the distribution of scores in control schools), $\bar{y}_{-icst}$ is the average contemporaneous score of her peers, X_{icst} is a vector of child-specific characteristics and ω_s, λ_t are school and year fixed effects, respectively. The repeated randomization of schools into treatment

[17]The evidence in Table 4.1 shows balance between the treatment and control groups in terms of observable characteristics.

and control groups over time allows for the inclusion of school fixed effects in equation (4.2). This allows us to control for time-invariant determinants of student achievement at the school level. We instrument $\bar{y}_{-icst}$ in (4.2) using the school treatment status.[18] Given the potential for error correlation across students within a given peer group, we cluster all standard errors at the class level. The coefficient ρ is the *endogenous* peer effect (Manski (1993)). This captures the effect of peers' contemporaneous test scores on individual achievement.[19]

Our identification strategy rests on the assumption that the treatment did not have any *direct* impact on the *non-treated*. This effectively means we have one variable that can be excluded from (4.2) while generating random variation in peers' average score. One potential concern is that by physically removing low-performing children from the classroom, high-performing students experienced a reduction in class size which *directly* affected their test scores. In Section 4.5, we discuss this and other potential threats to identification, and we perform several robustness checks to address them.

4.4 Results

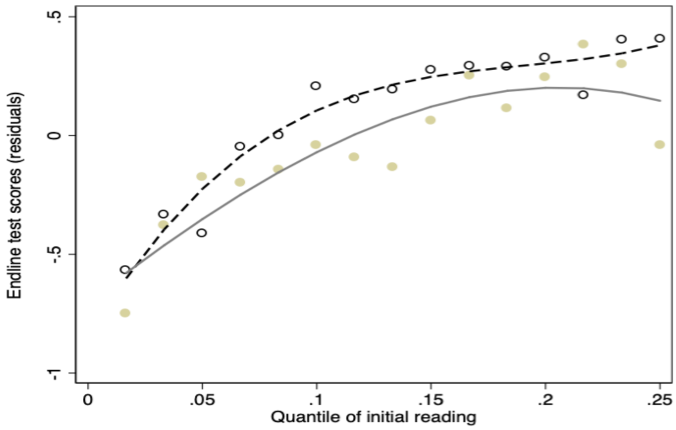
As shown in Section 4.2.4, the intervention generated experimentally induced variation in the outcomes of a subset of the students within the class. This result is the basis of our 2SLS strategy. We start by presenting graphical and regression-based evidence of the reduced-form effect of being in a treatment class on the sample high-achieving students in Section 4.4.1. Then, in Section 4.4.2 we estimate linear-in-means models of peer effects; we regress non-eligible students' test scores on the average *contemporaneous* score of their peers. We also explore whether these effects are heterogeneous depending on baseline achievement.

4.4.1 Reduced-form Evidence

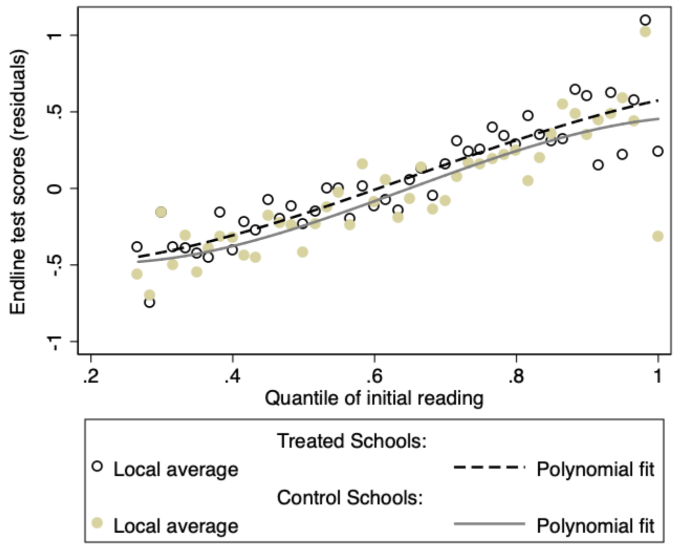
To assess the *indirect* effect of the remedial education program on high-achieving students, in Figure 4.2 we plot, separately for treated and control schools, students' end-line test scores as a function of baseline scores using a second-order polynomial. In each graph, we plot local averages and polynomial fits estimated separately for the treatment and control groups. For comparison purposes, we start by presenting the test scores of low-achieving students, those that were directly targeted by the intervention (see Panel A). Perhaps unsurprisingly given the results in Table 4.2, the fitted values in treatment schools are consistently above those in control schools for this sample of low achievers.

[18] Using the class treatment status does not make any difference because randomization took place at the school level. Therefore all classes within the same school experienced the same treatment.

[19] The existence of endogenous peer effects in the production function for test scores can be micro-founded using an effort game in the classroom, in which students' effort is determined jointly with peers' effort (see Fruehwirth (2013) and Tincani (2017)).



(a) Low-achieving Students



(b) High-achieving Students

Figure 4.2: Local Averages and Polynomial Fits of End-line Scores by Quantile of Baseline Reading

Dots represent local averages. Lines represent polynomial fits of end-line test scores as a function of baseline scores estimated using a second-order polynomial. The variable used to construct these figure is the residuals of standardized test scores obtained from a regression of end-line test scores on a second-order polynomial in age, gender and school fixed effects estimated separately for low-achieving students (Panel A) and high-achieving students (Panel B). By construction, the residuals are centered around zero.

Panel B illustrates the "reduced-form" effect of being in a treatment classroom for those students who were not eligible to receive the intervention because their baseline test score fell above the eligibility cutoff. More surprisingly, the same general picture observed for eligible children emerges for the non-eligibles, too. High-achieving students in treatment schools systematically outperform similar students in control schools. This is true for all quantiles of the baseline achievement distribution, even if there is some suggestive evidence that the effects are stronger at higher quantiles (we return to this point in Section 4.4.2). As we would expect, the magnitude of the difference in test scores between students in treatment and control schools is smaller for high-achieving students than for low-achieving students.

In Table 4.3 we present the reduced-form estimate of the impact of the intervention on high-achieving students. In column 1 we regress non-eligible students' outcomes on an indicator variable that takes the value of one if her school was in the treatment group, and zero otherwise. In columns 2 and 3 we include additional individual and class controls. Panel A reports the results for literacy, while panels B and C report those for math and total scores, respectively.

Consistent with the results shown in Figure 4.2, the literacy scores of high-achieving students in the treated schools were 0.108 standard deviations greater than the scores of similar students in the control schools (the p-value of the difference is 0.064). The effect is slightly larger when we control for individual characteristics (0.112 standard deviations, with a p-value of 0.053) and class characteristics (0.118 standard deviations, with p-value of 0.04). This coefficient is sizeable and represents roughly 30 percent of the treatment effect on eligible children. This result is economically meaningful, and its magnitude can be compared to a more commonly proposed school-level reform, tracking by prior achievement. Duflo, Dupas, & Kremer (2011) find that tracking raises test scores by 0.139 standard deviations both for students in upper and lower tracks.

Similarly to the results in Table 4.2, we find small and non statistically significant effects in math. This is reassuring and gives us confidence that these effects are indeed driven by peer-to-peer learning. We discuss this issue in more detail in Section 4.5. Because we do not find any reduced-form effect for math scores for non-eligible children, and given the results for treated students shown in Table 4.2, we focus on literacy and total scores.

4.4.2 Peer Effects in Academic Achievement

We now turn to the estimation of linear-in-means models of peer effects by estimating equation (4.2) on the sample of high-achieving students only. Table 4.4 report the OLS and 2SLS estimates. The OLS results in Panel A show that a one-standard-deviation increase in peers' contemporaneous achievement is correlated with an increase in literacy scores by 0.535 of a standard deviation (column 1). The result for total scores,

Table 4.3: Reduced-form Estimates: High-achieving Students

	(1)	(2)	(3)
Panel A: Literacy	0.108	0.112	0.118
	(0.058)	(0.058)	(0.057)
Panel B: Math	0.034	0.035	0.038
	(0.049)	(0.049)	(0.049)
Panel C: Total score	0.112	0.116	0.120
	(0.056)	(0.056)	(0.055)
Observations	5181	5181	5181
Individual controls		✓	✓
Class controls			✓

Notes: The outcome variables are standardized test scores. Individual controls include a second-order polynomial in age and gender. Class controls include class size and number of classrooms in the schools. All regressions control for school fixed effects. Robust standard errors are clustered at the classroom level, and presented in parentheses.

shown in column 4, implies that a one-standard-deviation increase in average peers' scores is associated with an increase in individual achievement by 0.56 of a standard deviation.

In panels B and C we report the first and second stage for the 2SLS model that uses the treatment status as an instrument for peers' average contemporaneous scores.[20] We have a very strong first stage: average peers' literacy score is 0.159 of a standard deviation higher in treatment classes compared to control classes (p-value = 0.007). By dividing the reduced-form coefficient (column 2 of panel A in Table 4.3) by the first-stage coefficient, we obtain an estimate of the peer effect parameter in equation (4.2). The 2SLS coefficient in column 1 implies that a one-standard-deviation increase in peers' contemporaneous achievement increases own achievement by 0.679 of a standard deviation. Column 4 reports the results for total scores, which are very similar, and imply that a one-standard-deviation increase in average peer end-line test scores would increase the test score of a student by 0.704 of a standard deviation.[21] These effects are comparable to those found in previous work. For instance, Boozer & Cacciola (2001) estimates an effect of 0.92 of a standard deviation for third-grade students, while Lavy & Schlosser (2011) finds a peer coefficient of 0.84. Using data from the Project STAR experiment, Whitmore (2005) finds that peers' score increases the individual score by 0.6 of a standard deviation.[22]

[20]The reduced form was reported in Table 4.3, and is therefore omitted here.

[21]As pointed out by Duflo, Dupas, & Kremer (2011), these results come from variation in peers' average achievement that are smaller than one standard deviation, so the extrapolation to one standard deviation might not be precise if the effects are non linear.

[22]Interestingly, as in De Giorgi, Pellizzari, & Redaelli (2010) who also estimate endogenous peer effects in the context of major choice, we find that the IV coefficient is larger than the OLS coefficient.

Because the previous literature has found evidence of non-linearities in peer effects (see Sacerdote (2001), Burke & Sass (2013), Tincani (2017)), we investigate whether the same is true for endogenous peer effects. To this aim we split the sample of non-eligible children using three terciles of the baseline achievement distribution, and estimate separate models for these three sub-samples. Table 4.5 reports the first-stage and second-stage regressions separately for students in the first, second, and third terciles of the baseline distribution of the outcome variable. For comparability purposes, in Panel A we report the second-stage coefficients from Table 4.4.

The results are consistent with the notion that students at the top of the achievement distribution benefit the most from improvements in their peers' outcomes.[23] In particular, we find that the peer effect coefficient monotonically increases with a student's baseline achievement quantile. Students just above the eligibility cutoff do not seem to be affected by their peers, while for students at the top of the distribution a one-standard-deviation increase in peers' contemporaneous score increases own literacy score by 0.777 of a standard deviation, and total score by 0.832 of a standard deviation (Panel C of Table 4.5).[24]

While we cannot estimate the effect on the lowest-ability students, as these were *directly* affected by the program, we find evidence that the endogenous peer effect is stronger for highest-ability student compared to "average"-ability students.[25]

[23]Interestingly, some of the papers that look at heterogeneous peer effects in academic achievement using the random allocation of students to peer groups find that high-ability students are less affected by peers' score than lower-ability students. For example, Carrell, Fullerton, & West (2009) find that the peer effect coefficient is larger for students in the bottom third of the academic ability distribution (even though they cannot reject the equality of the coefficients). A similar result is reported in Booij, Leuven, & Oosterbeek (2017). Importantly these papers only estimate a composite parameter that incorporates both the endogenous and exogenous peer effects.

[24]It is important to note that the standard errors are similar across subsamples, providing evidence that the insignificant effects on the lower quantiles stems from the low magnitude of the estimates, not from a lack of statistical power.

[25]We have also tried expanding equation (1) to allow the peer coefficient to vary with student's baseline achievement level and estimate: $y_{icst} = \rho_1 \bar{y}_{-icst} \times Q_1 + \rho_2 \bar{y}_{-icst} \times Q_2 + \rho_3 \bar{y}_{-icst} \times Q_3 + X_{icst}\beta + \omega_s + \lambda_t + \epsilon_{icst}$ Where Q_1, Q_2 and Q_3 are indicator variables taking the value one if child i falls in the first, second, or third tercile of the baseline achievement distribution. The results, shown in Appendix Table 4.8, follow the same patterns as those shown in Table 4.5, but they are somewhat larger for children in the third tercile of the baseline distribution. In that model, we test and reject the hypothesis that $\rho_1 = \rho_2 = \rho_3$.

Table 4.4: Linear-in-means Model of Peer Effects in Academic Achievement

	Literacy			Total		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - OLS						
Peers' average end-line score	0.535	0.533	0.527	0.560	0.559	0.555
	(0.048)	(0.048)	(0.049)	(0.046)	(0.046)	(0.046)
Panel B, 2SLS: first stage						
Treated class	0.159	0.159	0.168	0.159	0.159	0.166
	(0.059)	(0.059)	(0.058)	(0.056)	(0.056)	(0.055)
Panel C, 2SLS: second stage						
Peers' average end-line score	0.679	0.705	0.704	0.704	0.725	0.725
	(0.185)	(0.179)	(0.171)	(0.169)	(0.165)	(0.158)
Observations	5181	5181	5181	5181	5181	5181
Individual controls		✓	✓		✓	✓
Class controls			✓			✓

Notes: Estimates from models of dependent variable in column heading as a function of peers' average test scores. The outcome variables are standardized test scores. Individual controls include a second-order polynomial in age and gender. Class controls include class size and number of classrooms in the schools. All regressions control for school fixed effects. Robust standard errors are clustered at the classroom level, and presented in parentheses. Panel A reports the OLS results. Panel B and C report, respectively, the first-stage and second-stage results of a 2SLS model using random assignment to treatment as an instrument for peers' average end-line score.

Table 4.5: Heterogeneity by Baseline Achievement

	Literacy			Total		
Panel A: Linear-in-means model						
Peers' average end-line score		0.704			0.725	
		(0.171)			(0.158)	
Quantile of baseline achievement	Q1	Q2	Q3	Q1	Q2	Q3
Panel B, 2SLS: first stage						
Treated class	0.217	0.125	0.183	0.199	0.141	0.169
	(0.062)	(0.056)	(0.067)	(0.061)	(0.058)	(0.062)
Panel C, 2SLS: second stage						
Peers' average end-line score	0.277	0.539	0.777	0.339	0.546	0.832
	(0.265)	(0.458)	(0.374)	(0.283)	(0.361)	(0.376)
Observations	1726	1726	1724	1748	1715	1712

Notes: Estimates from models of dependent variable in column heading as a function of peers' average test scores. The outcome variables are standardized test scores. Controls include a second-order polynomial in age and gender, class size and number of classrooms in the schools. All regressions control for school fixed effects. Robust standard errors are clustered at the classroom level, and presented in parentheses. Panel A reports the second stage of the 2SLS model in columns 3 and 6 of Table 4.4. Panels B and C report the first and second stages of a 2SLS model using random assignment to treatment as an instrument for peers' average end-line score separately for children in the first, second, and third terciles of the baseline achievement distribution of the outcome variable.

4.5 Discussion

4.5.1 Threats to Identification and Mechanisms

As discussed in Section 4.3, our identification strategy rests on the assumption that the intervention does not directly effect learning outcomes of high-achieving students in treatment schools. If this was not the case there would not be any source of exogenous variation in average peers' scores that we could use to implement an instrumental variable strategy. While this assumption is not testable, in this section we do our best to rule out possible alternative mechanisms that could explain the increase in test scores that we observe for high-achieving students.

Class size

One potential concern is that by physically removing low-performing children from the classroom, high-performing students experienced a reduction in class size which had a *direct* positive impact on their test scores.

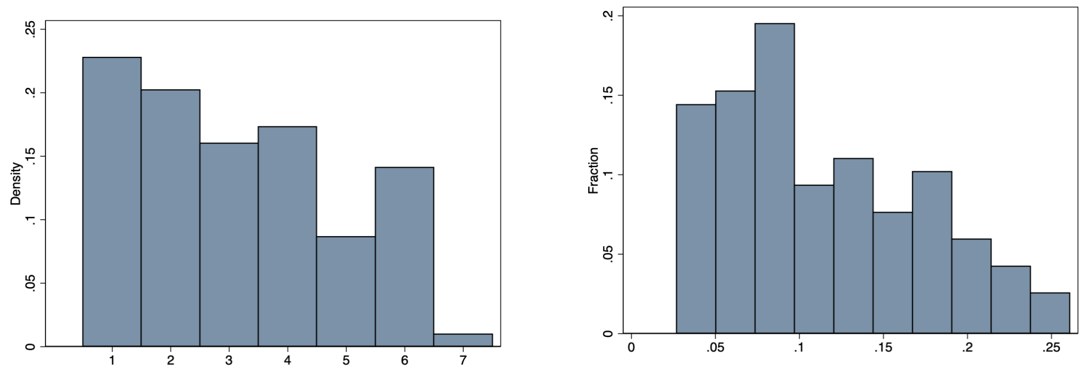
As we have discussed earlier, low-achieving students in treatment schools were randomized into different tutorial groups that took place at different times during the school day.[26] Figure 4.3 shows the distribution of same class students assigned to the same tutorial and the implied reduction in class size. In over 20 percent of classes only one student was assigned to the same tutoring group, and in more than 75 percent of classes fewer than four low-achieving students were randomized into the same group (Panel A).[27] The implied class size experienced by regular students, shown in Panel B, was composed of roughly three fewer students on average (out of an average class size of 31 students). Moreover, this reduction took place for only 40 minutes a day, three days a week, for a period of 16 weeks compared to the whole academic year (as in most studies on class size). This means that the class size reduction experienced by regular students lasted for roughly 32 hours out of almost 1,000 yearly school hours.

The paper that documents the largest class size effects in the literature is Urquiola (2006), which finds that reducing class size by on average nine students increases test scores by between 0.16 and 0.3 of a standard deviation.[28] Our reduction in class size is substantially smaller than the one reported in that paper, and lasted for a significantly shorter period of time. A back-of-the-envelope calculation implies that our reduction in class size predicts at most an increase in the test scores of regular students in the range of 0.005 to 0.01 of a standard deviation. This can explain at

[26] This was the case because there was a single tutor per school.

[27] Notice that while the size of the tutorial groups was capped at six, our records show that there is one school where the actual size was of the tutorial was increased to seven.

[28] Some of the results in Urquiola (2006) derive from up to three years of smaller class sizes. Moreover, the results reported in that study do not come from an RCT. The only study we are aware of that looks at the effect of reducing class size on learning outcomes in a developing country using an RCT is Duflo, Dupas, & Kremer (2015). This study finds no significant test gains for students exposed to a smaller class size.



(a) Number of Same Class Students per Group

(b) Class Size Reduction (%)

Figure 4.3: Peers in the Same Tutorial Group and Class Size Reduction (Treatment Schools Only)

Panel A shows the number of low-achieving students that attend the same class and were assigned to the same tutorial group. Panel B shows the class size reduction in our sample expressed in terms of total number of students in the classroom. The average class size reduction is three. In the sample 75 percent of classes experienced a reduction of 4.5 or fewer students.

most a tenth of the reduced-form effect found in Table 4.3, and is therefore unlikely to be driving our results.

An additional piece of evidence against the class size story comes from the lack of reduced-form effect on math test scores. The remedying tutorials did not take place specifically during regular literacy hours. Thus, if the reduction in class size were the main driver of our results, we would expect to see a reduced-form effect on this outcome as well. The fact that we did not find any economically meaningful and statistically significant effect on math scores for regular students in Table 4.3 (point estimate of 0.034 with an associated standard error of 0.049) further rules out large effects coming from a reduction in class size.

Teacher responses

The term peer effect is generally used as an umbrella term that comprises any externality implying that peers' outcomes affect an individual's outcome. This effectively includes peer-to-peer learning, student misbehaviour, and teacher practices (Sacerdote (2011)). Nonetheless, in this paper we seek to distinguish between these alternative mechanisms because such distinctions might be key for the design of optimal education policies.

While we are confident that the strategy presented earlier does not suffer from any of the identification issues described in Section 4.3, it does not allow us to disentangle the effects coming from student-to-student interactions from those that stem from teachers' behaviour. In particular, it might be that the treatment changed teachers' practices in ways that are key for student achievement. For example, by bringing new tutors to the schools, the treatment could crowd out teacher effort, *directly* affecting

Table 4.6: Linear-in-means Model of Peer Effects in Academic Achievement - Treatment Schools Only

	Literacy (1)	Total (2)
Panel A: OLS		
Peers' average end-line score	0.44	0.501
	(0.1)	(0.086)
Panel B, 2SLS: reduced form		
Share of eligible students	0.789	0.609
	(0.328)	(0.357)
Panel C, 2SLS: first stage		
Share of eligible students	0.77	0.485
	(0.251)	(0.262)
Panel D, 2SLS: second stage		
Peers' average end-line score	1.025	1.255
	(0.328)	(0.553)
Observations	2602	2602

Notes: The outcome variables are standardized test scores. Controls include a second-order polynomial in age, gender, peers' baseline average achievement and school fixed effects. Robust standard errors are clustered at the classroom level, and presented in parentheses. Panel A reports the OLS results. Panel B reports the reduced form. Panel C reports the first stage, and Panel D reports the second stage of a 2SLS model using the share of treated students as an instrument for average end-line test scores.

regular students' test scores.[29] To assess whether the effects we find are mediated by teacher responses we use two alternative strategies.

First, we note that within the same school different classes have different *shares* of low -achieving students, and over time the share of low-achieving students in a school varies.[30]

Therefore, in the group of treatment schools, we have variation in the class share of treated students. We exploit this source of variation to implement an instrumental variable strategy in *treatment schools only.* This strategy is similar to the one described earlier, but now the average score of peers ($\bar{y}_{-icst}$ in equation (4.2)) is instrumented using the share of eligible (hence, treated) students. Identification here is achieved using (i) idiosyncratic variation in the proportion of low achievers within a school over time, and (ii) between-class variation in the proportion of low-achieving students within the same school. We further control for average student achievement at baseline, so that we effectively compare classes that are similar in terms of average performance. By considering *treatment schools only*, we make sure that all teachers

[29] If this were the case, and if students' test scores were increasing in teacher effort, then the coefficient we are estimating would be downward biased.

[30] As we would expect in the absence of tracking, most of the within-school variation in our data comes from variation over time, rather than between classes in the same time period. In any given year, the between-school variation accounts for 66 to 83 percent of the total variation in the share of low-achieving students.

are being exposed to the same "treatment," so that the effects on non-eligible students cannot be explained by teaching practices that change discontinuously with the treatment (as in the earlier example).[31] Under the assumption that idiosyncratic variations in the share of low-achieving students (controlling for average baseline achievement and school fixed effect) within a school over time do not systematically affect teacher practices in ways that matter for regular students' test scores, this identification strategy allows us to tease out the peer effect coming exclusively from student-to-student interactions, net of any teacher response.

The results are presented in Table 4.6. Panel A shows the OLS results, while panels B and C report the reduced-form and first-stage results. Conditioning on school fixed effects and average class achievement at baseline, increasing the share of treated students by 10 percentage points increases the average peer score by 0.077 of a standard deviation (p-value = 0.003).[32] By dividing the reduced form by the first-stage coefficient, we calculate a peer-effect coefficient of 1.025 (Panel D of Table 4.6). This is not statistically different from the value of 0.705 found in Table 4.4. This allows us to rule out some particular types of teacher responses that could explain our results. Notably, any change in behaviour that occurs because of the treatment would be taken care of by this identification strategy.[33]

To further rule out effects stemming from teachers' responses, we present the results from a teacher survey that was administrated in a subsample of schools in our study sample. The survey included a set of items adapted from the teacher section of the Patterns of Adaptive Learning Scales (PALS). These scales are used to evaluate the teachers' outlook on the school goals, approaches to teaching, and teaching efficacy (Midgley et al. (2000)). We focus on three subscales of the PALS: (i) "Performance approaches" refers to the strategies used by teachers to convey to students that the purpose of engaging in academic work is to demonstrate competence. (ii) "Teacher

[31] A more subtle issue is that teachers practices might still be affected by the share of low-achieving students in the class. If this were the case, this second identification strategy would not allow us to control for teacher responses, and therefore, to identify the direct student-to-student spillover. While we do not think this is a very compelling story, the following example illustrates a scenario in which our strategy does not effectively control for teacher responses. If, controlling for school fixed effects and the average achievement of the class, a teacher were to change her behaviour when confronting a class that included a 10 percent share of low achievers as opposed to one including 30 percent of low achievers, then we would not be able to separately identify the effects of peers from those of teacher practices. Notice that by including school fixed effects we effectively control for differences in teaching strategies between schools. Finally, in the model proposed by Duflo, Dupas, & Kremer (2011), the relevant margin that matters for teachers' instructions is the median achievement level of the class. In our regressions we are controlling for mean achievement, so that teacher responses coming through that margin would be addressed. For this not to be the case, we would need teachers to act upon the share of low achievers, rather than mean achievement.

[32] As a falsification test, we estimated the same model in the sample of control schools as well. In the control group, we would expect a small and nonsignificant first stage because no remedial education intervention took place in these schools. We find this to be the case. A 10 percentage point increase in the share of low-achieving students translates in a non-significant increase of 0.007 of a standard deviation in the average end-line score of non-eligible students' peers (the results are available from the authors upon request).

[33] Interestingly, even if not statistically different, the point estimate in Table 4.6 is larger than that reported in Table 4.4, which would be consistent with the "crowding out" story described earlier.

Table 4.7: Teachers Reports

	Teaching Efficacy (1)	Performance Approaches (2)	Students' Bad Behavior (3)
Treated class	0.081	0.039	-0.367
	(0.273)	(0.215)	(0.293)
Observations	70	70	70

Notes: The outcome variables are factor scores constructed from the Patterns of Adaptive Learning Scales (PALS). These have been standardized to have a mean of zero and a standard deviation of one in the control group. Controls include teacher age, gender and experience. Standard errors are presented in parentheses.

efficacy" relates to teachers' beliefs that they are contributing significantly to the academic progress of their students, and that they can effectively teach to all students in their class. (iii) "Student bad behaviour" captures the extent to which teachers have to deal with student misconduct during school hours. Using the items from each of these subscales, we construct a composite using principal component analysis, which we then standardize to have a mean of zero and standard deviation of one in the control group.

To analyze whether there are differences in the behaviour of the teachers in treatment and control schools, we regress each outcome on a treatment indicator variable, controlling for teacher characteristics.[34] Table 4.7 shows the results. For all three outcomes we cannot reject the null hypothesis of equality between teachers in treatment and control schools. However, with the small sample size we have we might lack the statistical power needed to detect any significant differences. Interestingly, we find that the point estimates in columns 1 and 2 are very small in magnitude, while the point estimate from column 3 is substantially larger. This provides some suggestive evidence that teachers reported having to deal with students' misbehaviour more often in control schools than in treatment schools. This result is consistent with previous work showing that classroom disruptions increase with the level of low-ability students (see Carrell & Hoekstra (2010) and Lavy, Paserman, & Schlosser (2011)). On the other hand, the fact that teacher practices do not seem to be affected by the intervention is at odds with the evidence in Duflo, Dupas, & Kremer (2011). However, it is important to keep in mind that the intervention considered in this paper did not involve teachers in any way. Nor did any direct contact between teachers and tutors occur. In particular, there was no change in the student body composition as teachers kept teaching to the exact same students throughout the experiment.

[34]The results when we do not include teacher characteristics are virtually identical and are not reported.

4.5.2 Implications

The findings in this paper imply that failing to consider the indirect effects of the remedying intervention on non-eligible students would underestimate the true treatment effect for the overall student population. Consider the following back-of-the-envelope calculation: The intervention cost USD 89 per eligible student in 2016 (Alvarez Marinelli, Berlinski, & Busso (2019)). Using our results from Table 4.2, we calculate that this implies that for every USD 100 spent, low-achieving students' test scores increased by 0.406 of a standard deviation. In addition, for every USD 100 spent, high achievers' test scores also increased by 0.121 of a standard deviation.[35] Given that there are three times more non-eligible students than eligible students, this translates into an additional increase of 0.363 of a standard deviation in test scores for every USD 100 spent. Therefore, a failure to consider the treatment effect on high achievers would result in a 47 percent underestimate of the treatment effect on test scores in the population of students in treatment schools.[36]

The results in this paper thus underline the need to collect data on the entire local economy to fully appreciate policy effects and correctly compute the *returns* to remedial education policies. Endogenous peer effects lead to a social-multiplier effect that amplifies the total output of a program. Finally, from a methodological point of view, our findings emphasize the importance of experimentally manipulating individuals' treatment status within treatment units (schools in our setting) to identify social interactions.

4.6 Conclusions

In this paper, we report the results from a randomized evaluation targeting children at the bottom of the class. Wed show that the intervention was effective in improving everyone's learning: both low achievers and high achievers' academic outcomes were substantially higher in treatment schools compared to control schools. The intervention we consider is a remedying education program aimed at improving reading among struggling third-grade students in Colombia. In treatment schools, students with baseline reading score lower than a certain threshold were assigned to small group tutoring classes during which they worked with a qualified tutor following a structured pedagogical curriculum. The intervention was very effective in improving literacy skills in the sample of low-achieving students: average test score increased by 0.362 of a standard deviation by the end of the intervention.

Importantly, the randomization strategy naturally generates two groups of students within the same class: a group of low-achieving students who were *eligible* to receive the intervention, and a group of higher-achieving students who were *not eligible*. We

[35] For eligible students, this is given by $0.362 \times (\frac{100}{89}) = 0.407$. For non-eligible students, using our results from Table 4.3 this is given by $0.108 \times (\frac{100}{89}) = 0.121$.

[36] This is given by $\frac{0.363}{(0.363+0.407)} = 0.47$.

can therefore study whether an exogenous change in the test scores at the bottom of the class translates into gains at the top. We find that this is the case. *Non-eligible* children in treated schools scored 0.108 of a standard deviation higher than similar children in the control group. This coefficient is sizable and represents roughly 30 percent of the treatment effect on the *eligible* students.

Using the treatment-induced variation in peers' scores as an instrument for peers' outcomes, we estimate a linear-in-means model of peer effects, focusing particularly on their endogenous component. The random allocation of the treatment allows us to overcome the identification challenges that have plagued much of the previous literature on peer effects, namely selection, reflection, and correlated unobservables (Manski (1993)). We find strong evidence of peer effects in academic outcomes: a one-standard-deviation increase in peers' *contemporaneous* test scores increases individual reading scores by 0.679 of a standard deviation. We find evidence of non-linearities, with largest effects at the top of the ability distribution. We further rule out alternative mechanisms coming from a reduction in class size, or changes teacher practices. We find some suggestive evidence that some of the effect might be due to a reduction in students' misbehaviour.

This study provides the first successful example of how peer effects can be exploited in the design of public policies aimed at improving students' academic performance. Taken together, our findings suggest that policies aimed at improving the bottom of the achievement distribution have the potential to generate social-multiplier effects. This indicates that it is possible to substantially improve the quality of education *for all* with relatively cheap and easy-to-scale interventions. The findings provide a strong rationale that underscores why society should care about improving the educational outcomes of the weakest. These considerations are important to inform any policy debate concerned with the allocation of public funds to education.

Appendix 4.A Appendix Tables

Table 4.8: Heterogeneity by Baseline Achievement - Interacted Model

	Literacy			Total		
	Q1	Q2	Q3	Q1	Q2	Q3
Panel A: first stage						
Treated class	0.655	0.680	0.754	0.527	0.546	0.621
	(0.044)	(0.049)	(.062)	(0.042)	(0.035)	(0.058)
Panel B: Second stage						
Peers' average endline score	0.069	0.614	1.148	-0.301	0.485	1.119
	(0.253)	(0.252)	(0.245)	(0.296)	(0.293)	(0.276)
Panel C: F-test of joint equality						
$(\rho_1 = \rho_2 = \rho_3)$						
p-value:		0.000			0.000	
Observations		5181			5181	

Notes: Estimates from models of dependent variable in column heading as a function of peers' average test scores. The outcome variables are standardized test scores.Controls include a second order polynomial in age and gender, class size and number of classrooms in the schools. All regressions control for school fixed effects. Robust standard errors are clustered at the classroom level and presented in parenthesis. Panels A and B report the first and second stages of a 2SLS model using classroom random assignment to treatment interacted with student quantiles as an instrument for peers' average endline score.

References

Adams, A., & Andrew, A. (2019). *Preferences and beliefs in the marriage market for young brides* (Tech. Rep.). IFS Working Papers.

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

Agostinelli, F., & Wiswall, M. (2016a). *Estimating the technology of childrens skill formation.* (NBER Working Paper No. 22442)

Agostinelli, F., & Wiswall, M. (2016b). *Identification of dynamic latent factor models: The implications of re-normalization in a model of child development.* (NBER Working Paper No. 22441)

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

Alcadía de Manizales. (2017). Boletín estadístico del sector educativo año 2017, Manizales, Caldas.

Almond, D., & Mazumder, B. (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, *3*(4), 56–85.

Alvarez Marinelli, H., Berlinski, S., & Busso, M. (2019). Remedial reading in small group tutorials. short and medium run effects of three experiments in colombia. *Unpublished manuscript*.

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.

Anderson, T., & Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the third berkeley symposium on mathematical statistics and probability* (Vol. 5, p. 111-150). Berkeley: University of California Press.

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*.

Angelucci, M., De Giorgi, G., Rangel, M. A., & Rasul, I. (2010). Family networks and school enrolment: Evidence from a randomized social experiment. *Journal of public Economics*, *94*(3-4), 197–221.

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.

Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? evidence from boston's metco program. *American Economic Review*, *94*(5), 1613–1634.

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., 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., Blundell, R., Conti, G., & Mason, G. (2020). Inequality in socio-emotional skills: A cross-cohort comparison. *Journal of Public Economics*, *191*, 104171.

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.

Attanasio, O., Meghir, C., Nix, E., & Salvati, F. (2017). Human capital growth and poverty: Evidence from ethiopia and peru. *Review of Economic Dynamics*, *11*(8).

Baland, J.-M., & Robinson, J. A. (2000). Is child labor inefficient? *Journal of Political Economy*, *108*(4), 663–679.

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.

Bayley, N. (2005). Bayley scales of infant development.

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

Behrman, J. R., & Rosenzweig, M. R. (2004b). Returns to birthweight. *Review of Economics and Statistics*, *86*(2), 586–601.

Bernal, R., Giannola, M., & Nores, M. (2020). *The effects of a project and play-based early education program on medium term developmental trajectories of young children in a low-income setting.* (Unpublished manuscript Universidad de los Andes and National Institute for Early Education Research.)

Bernal, R., Martínez, M. A., & Quintero, C. (2015). *Situación de niñas y niños colombianos menores de cinco años, entre 2010 y 2013* (Tech. Rep.). Bogota, Colombia: Editorial Kimpress.

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

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

Bobonis, G. J., & Finan, F. (2009). Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics*, *91*(4), 695–716.

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.

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *The Review of Economic Studies*, *84*(2), 547–578.

Boonk, L., Gijselaers, H. J., Ritzen, H., & Brand-Gruwel, S. (2018). A review of the relationship between parental involvement indicators and academic achievement. *Educational Research Review*, *24*, 10-30.

Boozer, M., & Cacciola, S. E. (2001). Inside the'black box'of project star: Estimation of peer effects using experimental data.

Bracken, B. A., Keith, L. K., & Walker, K. C. (1998). Assessment of preschool behavior and social-emotional functioning: A review of thirteen third-party instruments. *Journal of Psychoeducational Assessment*, *16*(2), 153–169.

Brazelton, T. B. (1984). *Neonatal behavioral assessment scale.* (2nd ed. ed.). London: Blackwell Scientific/Spastics International.

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, intrahousehold 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.

Bullowa, M. (1979). *Before speech: The beginning of interpersonal communication.* CUP Archive.

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, *31*(1), 51–82.

Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, *82*(4), 1273–1301.

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.

Card, D., & Giuliano, L. (2016). Can tracking raise the test scores of high-ability minority students? *American Economic Review*, *106*(10), 2783–2816.

Carrell, S., Fullerton, R., & West, J. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, *27*(3), 439–464.

Carrell, S., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, *108*(11), 3377–3415.

Carrell, S., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, *2*(1), 211–28.

Carrell, S., Sacerdote, B., & West, J. (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, *81*(3), 855–882.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, *329*(5996), 1194–1197.

Channon, A. A. (2011). Can mothers judge the size of their newborn? Assessing the determinants of a mother's perception of a baby's size at birth. *Journal of biosocial science*, *43*(5), 555–573.

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.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 28–45.

Cipollone, P., & Rosolia, A. (2007). Social interactions in high school: Lessons from an earthquake. *American Economic Review*, *97*(3), 948–965.

Coleman, J. S. (1968). Equality of educational opportunity. *Integrated Education*, *6*(5), 19–28.

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

Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, *74*(5), 1191–1230.

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.

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

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

Currie, J., & Hyson, R. (1999). Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American Economic Review*, *89*(2), 245–250.

Currie, J., & Vogl, T. (2013). Early-life health and adult circumstance in developing countries. *Annu. Rev. Econ.*, *5*(1), 1–36.

Dahl, G. B., Løken, K. V., & Mogstad, M. (2014). Peer effects in program participation. *American Economic Review*, *104*(7), 2049–74.

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.

DeBruin-Parecki, A. (2005). *Early literacy skills assessment user guide.* Ypsilanti, MI: High/Scope Press.

De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, *2*(2), 241–75.

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.

Departamento Administrativo Nacional de Estadstica. (2018).

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.

Doyle, A., Ostrander, R., Skare, S., Crosby, R. D., & August, G. J. (1997). Convergent and criterion-related validity of the behavior assessment system for children-parent rating scale. *Journal of Clinical Child Psychology*, *26*(3), 276–284.

Dubeck, M. M., & Gove, A. (2015). The early grade reading assessment (egra): Its theoretical foundation, purpose, and limitations. *International Journal of Educational Development*, *40*, 315–322.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, *101*(5), 1739–74.

Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from kenyan primary schools. *Journal of Public Economics*, *123*, 92–110.

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of economics*, *118*(3), 815–842.

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.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). Tvip: Test de vocabulario en imgenes peabody: adaptación hispanoamericana. *American Guidance Service*.

Fernald, Prado, E., Kariger, P., & Raikes, A. (2017). A toolkit for measuring early childhood development in low and middle-income countries.

Figlio, D., Guryan, J., Karbownik, K., & Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, *104*(12), 3921–55.

Figlio, D. N. (2007). Boys named sue: Disruptive children and their peers. *Education Finance and Policy*, *2*(4), 376-394.

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.

Frongillo, E., Sywulka, S., & Kariger, P. (2003). *UNICEF psychosocial care indicators project. Final report to UNICEF.* (Mimeo, Cornell University)

Fruehwirth, J. C. (2013). Identifying peer achievement spillovers: Implications for desegregation and the achievement gap. *Quantitative Economics*, *4*(1), 85–124.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeerch, C., Walker, S., ... Grantham-McGregor, S. (2014, May). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, *344*, 998-1001.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., ... others (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet*, *369*(9555), 60–70.

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.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, *18*(5), 527–544.

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

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

Heckman, J., Moon, S., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of Public Economics*, *94*, 114-128.

Hesselius, P., Nilsson, J. P., & Johansson, P. (2009). Sick of your colleagues'absence? *Journal of the European Economic Association*, *7*(2-3), 583–594.

Hoxby. (2000). *Peer effects in the classroom: Learning from gender and race variation* (Tech. Rep.). National Bureau of Economic Research.

Hoxby, & Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects* (Tech. Rep.). Working paper.

Hsin, A. (2007). Mothers time with children and the social reproduction of cognitive skills..

Hsin, A. (2012). Is biology destiny? birth weight and differential parental treatment. *Demography*, *49*(4), 1385–1405.

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*.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, *102*(5), 2048–82.

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.

Johnson, R. C., & Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from head start and public school spending. *American Economic Journal: Economic Policy*, *11*(4), 310–49.

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.

Lalive, R., & Cattaneo, M. A. (2009). Social interactions and schooling decisions. *The Review of Economics and Statistics*, *91*(3), 457–477.

Lavy, V., Paserman, M. D., & Schlosser, A. (2011). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, *122*(559), 208–237.

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*(2), 1–33.

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, *116*(3), 777–803.

Lefgren, L. (2004). Educational peer effects and the chicago public schools. *Journal of Urban Economics*, *56*(2), 169–191.

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

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

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *The Review of Economics and Statistics*, *89*(2), 289–299.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542.

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.

McCormick, M. P., Ketner Weissman, A., Weiland, C., Hsueh, J., Sachs, J., & Snow, C. (2020). Time well spent: Home learning activities and gains in childrens academic skills in the prekindergarten year. *Developmental Psychology*.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, *27*(1), 415–444.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., ... others (2000). Manual for the patterns of adaptive learning scales. *Ann Arbor: University of Michigan*.

Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, *4*(45-82), 6–17.

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.

Muñoz-Sandoval, A., Woodcock, R., McGrew, K., & Mather, N. (2005). Batería iii pruebas de aprovechamiento. *Itasca, IL: Riverside Publishing*.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019, April). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review*, *109*(4), 1426-60.

Nores, M., Bernal, R., & Barnett, S. (2019). Center-based care for infants and toddlers: the aeiotu randomized trial. *Economics of Education Review*, *72*, 30 - 43.

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

Padilla, C. M., & Ryan, R. M. (2018). School readiness among children of hispanic immigrants and their peers: The role of parental cognitive stimulation and early care and education. *Early Childhood Research Quarterly*.

Padilla, E., Lugo, D., & Dunn, L. (1986). Test de vocabulario en imágenes peabody (tvip). *Circle Pines, MN: American Guidance Service (AGS), Inc.*.

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.

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental psychology*, *45*(3), 605.

Ponitz Cameron, C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*(2), 141–158.

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

Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, *92*(5), 1394 - 1415.

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

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.

Royer, H. (2009). Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, *1*(1), 49–85.

RTI-International. (2009). Early grade reading assessment toolkit. *World Bank Working Paper, Office of Human Development*.

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

Rubio-Codina, M., Araujo, C., Attanasio, O., Munoz, P., & McGregor, S. G. (2017). Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies. *PLOS-ONE*, *25*, 234-259.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of economics*, *116*(2), 681–704.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the economics of education* (Vol. 3, pp. 249–277). Elsevier.

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.

Sparrow, S. S., Cicchetti, D., & Balla, D. A. (1984). Vineland adaptive behavior scales.

Squires, J., Bricker, D., & Twombly, E. (2002). *The asq: Se user's guide: For the ages & stages questionnaires: Social-emotional.* Paul H Brookes Publishing.

Stinebrickner, R., & Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, *90*(8-9), 1435–1454.

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.

Susman-Stillman, A., Kalkoske, M., Egeland, B., & Waldman, I. (1996). Infant temperament and maternal sensitivity as predictors of attachment security. *Infant Behavior and Development*, *19*(1), 33–47.

Tincani, M. (2017). Heterogeneous peer effects and rank concerns: Theory and evidence, CESifo Working Paper Series.

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

United Nations. (2016). Sustainable development goals. goal 4: Ensure inclusive and quality education for all and promote lifelong learning. *http://www.un.org/sustainabledevelopment/education/*.

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

Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural bolivia. *The Review of Economics and Statistics*, *88*(1), 171–177.

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. *American Economic Review*, *95*(2), 199–203.

WHO. (2007). *Who child growth standards: head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: methods and development*. World Health Organization.

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.

World Bank. (2018). World development report 2018: Learning to realize education's promise. *Washington, DC: World Bank*.

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*.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, *85*(1), 9–23.

Statement of Conjoint Work

Note on the joint work in Michele Giannola's thesis "Essays on the Economics of Human Capital".

The chapter, "Child Development in the Early Years: Parental Investment and the Changing Dynamics of Different Dimensions" is co-authored work between Michele Giannola, Orazio Attanasio, Raquel Bernal and Milagros Nores and each author contributed equally.

The chapter, "Parental Investments and Intra-household Inequality in Child Development: Theory, Measurement, and Evidence from a Lab-in-the-Field Experiment" , is single authored by Michele Giannola.

The chapter,"Helping Struggling Students and Benefiting All: Peer Effects in Primary Education" is co-authored work between Michele Giannola, Samuel Berlinski and Matias Busso and each author contributed equally.