Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ECONOMICS

Ciclo 31

Settore Concorsuale: 13/A1

Settore Scientifico Disciplinare: SECS-P/01

Essays on Human Capital Development and Socio-economic Inequality

Presentata da: Anh Nguyet Tran Thi

Coordinatore Dottorato

Prof.ssa Maria Bigoni

Supervisore

Prof.ssa Chiara Monfardini

Esame finale anno 2019

Declaration

I certify that the thesis I have presented for examination for the PhD degree in Economics at the University of Bologna is my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). Chapter 1 and Chapter 2 are solely my work and Chapter 3 is jointly co-authored with Professor Chiara Monfardini (University of Bologna). The copyright of this thesis rests with the authors. Quotation from it is permitted, provided that full acknowledgement is made.

This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

Statement of Joint Work

I confirm that Chapter 3 is jointly co-authored with Professor Chiara Monfardini (University of Bologna), all authors contributed equally to the project.

Introduction

This dissertation consists of three chapters in which I address central research questions about the role of parental investments and family structure on human capital development, the impact of education on labour-market outcomes and learning outcomes, and the origins and mechanisms of inter-generational mobility in developing countries.

The first chapter examines how parental monetary investment affects the joint evolution of child health, cognitive skills and socio-emotional skills. I estimate a dynamic factor model, characterizing the skill formation process over the childhood, from birth to 12 years of age, using the sample of Vietnamese children from the Young Lives study. To deal with the endogeneity of parental investment caused by the parental response to unobserved heterogeneity, I take advantage of the plausibly exogenous variation in investment incentives generated by the geographic- and time-variant welfare rules under the 1999-2020 National Target Programs. The main findings are as follows. First, parental monetary investment is productive in promoting child capabilities during the childhood. However, the productivity of investments varies significantly across different periods and capabilities. Second, parental investments are more productive for children with higher initial stocks of health, cognitive skills, and socio-emotional skills. The presence of this complementarity is evident throughout the childhood for the health and cognitive development process. Third, the dynamic model reveals that child cognitive skills and health are highly and increasingly self-productive over time. Moreover, there is cross-productivity among different dimensions of child capabilities which corroborates the importance of their joint modelling. Lastly, children living in deprived conditions are not only disadvantaged by potentially receiving suboptimal parental investments, but also the environment exerts direct detrimental effects on their capabilities.

In the second chapter, I estimate marginal returns to upper secondary school on the labour market and on learning outcomes in Indonesia. Using the longitudinal data from the Indonesian Family Life Survey 1997-2015, I document a substantial degree of heterogeneity in the returns to upper secondary school on the labour market. Marginal returns in earnings are found to be higher for individuals with characteristics that make them more likely to attend upper secondary school. In contrary, students with higher gains on learning outcomes are less likely to attend school. Moreover, students from disadvantaged backgrounds are not only less likely to go to upper secondary school andbut also have substantially lower marginal returns on the labour market. These findings suggests that universal upper secondary school expansion that successfully attract low-resistant students who are currently not in upper secondary school may yield large pecuniary returns but are inequitable. Marginal expansions targeting disadvantaged students are likely to be both efficient and equitable than universal upper secondary school policies.

The third chapter investigates the origins and mechanisms of birth order effects on cognitive skills, socio-emotional skills and health in Vietnam. Using a sample of children from the Young Lives study we find strong evidence of negative birth order effects on parental investments and child capabilities, emerging very early in life. Parents invest significantly less money on second-born children from age one until age eight. Second-born children are less healthy, have lower stocks of cognitive skills and socio-emotional skills than the first borns. We go beyond the previous studies and decompose the long-term effects of birth order on child capabilities, showing the existence of three channels: (i) the production technology efficiency, (ii) the parental monetary investments, (iii) the self-productiveness of child capabilities. While the first mechanism is only significant in health production technology during adolescence, the contributions of the second and third mechanisms are sizeable for the three analysed capabilities at different stages of development.

Acknowledgement

I am grateful to my supervisor Chiara Monfardini for her support and advice throughout my PhD. I owe her a greater debt of gratitude than can be expressed. To experience Monfardini's openminded and positive interest in my ideas was heart-warming. She is extremely patient to listen to my problems, not always very well-formulated, then re-interpret them in a clear economic language. I am grateful for her habit to comment on every ambiguous explanation with an "Explain more". It certainly forced me to sharpen and refine the arguments as well as is invaluable for my understanding of puzzling features in the studies. In particular, I thank to her for having read and correct numerous imprecise interpretations, mistakes and typos on every line of the thesis. No doubt, Professor Monfardini has taught me to be a better economic graduate and a more economics enthusiast.

I am grateful to Paul Glewwe and Matthew Wiswall for generously providing insightful suggestions. The comments by Glewwe in the first round of thesis submission has helped me redefining the first chapter. Special thanks to Matthew John Wakefield, Margherita Fort, Maria Bigoni, Davide Dragone, and Daniela Iorio for their generosity and invaluable feedback.

I am grateful to Pedro Carneiro and Sonya Krutikova for hosting me during my visit to University College London. Special thanks to Carneiro for invaluable feedback on the second of this thesis. I have learnt innumerable lessons from Krutikova and other colleagues during the time in the RISE programme at the Institute of Fiscal Studies.

I am indebted to Christopher Gilbert who has been my mentor for seven years. The very first class of Gilbert I took at the University of Trento and his long-standing encouragement and advice has been always a source of inspiration. Without your patience and invaluable supports, my study in Bologna would have never been realized.

I want to thank my friends who always stand by me, sometimes listen to my ideas, sometimes exchange opinions, but often just give me strength and encouragement to carry on. In particular, I owe Anatasia Arabadzhya a big and tight hug for being such a great friend, for four years we have lived and studied together and for making my life in Bologna so much better. I am thankful to Yanjun Li, Ankush Asri, Anh Nguyen, Lien Nguyen for giving me strength and helping me become more optimistic and work harder.

I am thankful to my family: my parents for raising me, supporting me and standing by me at every step. To my father who has passed away during my third year of the degree: words cannot express my love and gratitude for you, I am sure you are proud of me wherever you are.

Contents

List of Figures

List of Tables

Chapter 1

Parental Monetary Investment and Child Capability Formation

Abstract

This paper examines how parental monetary investment affects the joint evolution of child health, cognitive skills and socio-emotional skills. I estimate a dynamic factor model, characterizing the skill formation process during childhood, from birth to 12 years of age, using the sample of Vietnamese children from the Young Lives study. To deal with the endogeneity of parental monetary investment caused by parental responses to unobserved heterogeneity, I take advantage of the plausibly exogenous variation in investment incentives generated by the geographic- and time-varying welfare rules under the 1999-2020 National Target Programs. The main findings are as follows. First, parental monetary investment is productive in promoting child capabilities during the childhood. However, the productivity of investments varies significantly across different periods and capabilities. Second, parental monetary investments are more productive for children with higher initial stocks of health, cognitive skills and socio-emotional skills. The presence of this complementarity is evident throughout the childhood for the health and cognitive development. Third, the dynamic model reveals that child cognitive skills and health are highly and increasingly self-productive over time. Moreover, there is cross-productivity among different dimensions of child capabilities which corroborates the importance of their joint modelling. Lastly, children living in deprived conditions are not only disadvantaged by potentially receiving suboptimal parental monetary investments, but also the environment they live exerts direct detrimental effects on their capabilities.

1.1 Introduction

A large body of literature on human capital development has established that individual capabilities are multi-dimensional in nature, including health, cognitive skills and socio-emotional skills, which altogether are important in determining later life outcomes (e.g., Cunha et al., 2006; Heckman, 2006; Heckman and Mosso, 2014; Attanasio et al., 2017; Agostinelli and Wiswall, 2018). Seminal studies by Cunha and Heckman (2008) and Cunha et al. (2010), using US data, have shown that the skill formation process consists of multiple stages, and more importantly, the productivity of parental monetary investments vary significantly over time.

Much prior empirical research has focused on the formation process of child cognitive skills and socio-emotional skills and the effects of parental inputs on the development of these skills. Though health is considered a crucial element of human capital, and investing on it provides enormous payoffs, the joint determination of health, cognitive skills, and socio-emotional skills, is not yet well understood. In this paper, I investigate the joint development of health, cognitive skills and socio-emotional skills using the conceptual framework of Cunha et al. (2006), Cunha and Heckman (2007), and Heckman and Mosso (2014). Specifically, I estimate production technologies of child health, cognitive skills and socio-emotional skills that includes as inputs parental monetary investments and family background factors over different stages of childhood. For the empirical analysis I use a sample of Vietnamese children from the Young Lives study. While existing evidence mostly refers to the developed economies, this study focuses on Vietnam and therefore adds to the body of small but growing literature on human capital development in developing countries. The Young Lives study traces the development process of the same children throughout their childhood, making available substantial information on child's capability measures, child-specific parental monetary investments, home environment and parental background factors.

The paper addresses the questions of *when* and *how much* to invest in children, and the interrelationship of different dimensions of human capital by three steps. The first step identifies and estimates the distribution of child health, cognitive skills, and socio-emotional skills over the childhood. I acknowledge that the observed measures are imperfect proxies of the underlying latent capability and quantify the extent of measurement errors. In estimating the latent factors, I do not impose any distributional assumptions on the latent factors and measurement errors. In the second step, I estimate a dynamic production function of child capabilities using the estimated latent factors from the first step. The production technology allows the productivity of parental monetary investment to vary across different stages and for the possibility of self- and cross-productivity between capabilities. That is, the existing stocks of child capabilities and parental monetary investments produce stocks of child capabilities in the next stage. I acknowledge the arbitrariness of the metrics of the latent factors and anchor child capabilities to adult outcomes (educational aspiration).

The first step addresses the conventional issue of measurement errors in measuring child capabilities. It has been well established in the literature that the use of imperfect proxies for child capabilities would produce biased estimates. Indeed, if the measurement errors are non-classical, it is difficult to sign the direction of estimate bias and to defend the standard instrumental strategy[1]. In this paper, I exploit the availability of multiple measures of health, cognitive skills, and socio-emotional skills in the Young Lives study over 12 years of childhood to identify and estimate the latent child capabilities which are not contaminated by measurement errors.

However, correcting for measurement errors is still not sufficient for the estimation of investment productivity if parents make investment decision with some knowledge about production shocks affecting child developments and this information is not available to the analyst. To secure the identification of investment productivity, I take advantage of the plausibly exogenous variation in investment incentives generated by the geographic- and time-variant welfare rules under the National Target Programme (NTP) implemented from 1999 to 2020 in Vietnam. Specifically, I exploit the potential changes in family wealth and/or income induced by household's eligibility to participate into the Housing Support Program, which is a specific component under the umbrella of the NTP. The HSP and more generally the NTP have been shown to improve the lives of participating families and communities, including their children. Eligible families would be entitled to the maximum benefits of: (i) an amount of conditional cash transfers equivalent to about 150% of income at poverty line; and (ii) a similar amount of unconditional low- or zero-interest loan.

The study reveals the following main findings, which contribute to the understanding of the development of child health, cognitive skills, and socio-emotional skills, emphasizing the role exerted by families through parental monetary investments, parental backgrounds and long-term financial resources in developing countries. First, parental monetary investments are effective in producing child capabilities from age 1 to age 12, which is divided into four periods: the initial period of age 0-1; the second period of age 2-5; the third period of age 6-8; and the fourth period from age 9-12. This investment productivity is decreasing over time for cognitive skills and socio-emotional skills, but is U-shaped with respect to health, that is investment in health is most productive in the earliest periods until age 5, and in the final period, from age 9 to age 12. Second, I find that the productivity of investments is higher for children with higher initial stocks of health, cognitive skills, and socio-emotional skills. The presence of this complementarity is evident throughout childhood for

[1] See Hausman (2001), Hausman, Newey and Powell (1995), Griliches and Ringstad (1970) and Schennach (2013). Specifically, for multivariate linear regressions and nonlinear specifications, when the measurement errors are nonclassical, the widespread folkflore that the estimated coefficients are downward biased fails to hold. The resulting statistical inference is also no longer valid. Moreover, the standard IV is also not applicable (Amemiya, 1985).

the health and cognitive development processes. Complementarity of investments and capabilities constitutes a fundamental source of human capital inequality among Vietnamese children.

Third, I find that child health and cognitive skills exhibit a strong pattern of self-productivity throughout the childhood which becomes stronger as children grow up. There is cross-productivity between capabilities - better health improves child cognitive stocks and cognitive skills promote socio-emotional skills. In particular, the cross-productivity of health on cognitive development and of noncognitive skills on health are significant and robust to different specifications of the production function. These findings emphasize the importance of modelling jointly the development of child health, cognitive skills and socio-emotional skills.

Fourth, by investigating parental monetary investments made to a specific child while simultaneously control for family wealth, I show the effects of socioeconomic status (SES) on the development of child health, cognitive skills, and socio-emotional skills. Children living in deprived conditions are not only disadvantaged by potentially receiving suboptimal parental monetary investments but also by the direct detrimental effects of the environment on their capabilities.

This paper builds on a number of earlier papers that estimate dynamic production technology of child capability in the United States, for example, Cunha and Heckman (2008), Cunha et al. (2010), Del Boca et al. (2014), Agostinelli and Wiswall (2016a, 2016b); and in the developing countries using the Young Lives study, such as Helmers and Patnam (2011), Attanasio et al. (2017a, 2017b, 2018). More recently, Agostinelli and Wiswall (2016a, 2016b) discuss a number of estimation issues relating to the latent factor approach, which is relevant to my work.

I depart from the aforementioned studies and contribute to the literature in three ways. Specifically, I investigate the joint evolution of health, cognitive skills, and socio-emotional skills over childhood while allowing for a greater flexibility of the production technologies, which exhibit Hicks neutral TFP dynamics and nonconstant returns to scale. Moreover, I account for the endogeneity of parental monetary investments by taking advantage of the plausibly exogenous variation in investment incentives generated by the geographic- and time-variant welfare rules. Cunha, Heckman, and Schennach (2010) develop the dynamic latent factor approach I use in this paper. They use this approach to estimate the process of cognitive and noncognitive skill accumulation over two stages of childhood for children in the U.S. aged 0-14, using the National Longitudinal Survey of Youth data. In their empirical analysis, they neither model health development nor allow for the presence of total factor productivity (TFP) and nonconstant returns to scale, which I do in this paper. I also allow for the possibility of complementarity (substitution) between investments and capability, and for the production technologies to exhibit Hicks-neutral total factor productivity (TFP) dynamics and non-constant returns to scale. Except for Agostinelli and Wiswall (2016a, 2016b), all of the aforementioned studies restrict the production function to exhibit constant returns to scale. However, Agostinelli and Wiswall (2016a, 2016b) estimate only the technology of cognitive skills production.

My study is also related to that of Del Boca, Flinn, and Wiswall (2014), which uses the U.S Panel Study of Income Dynamics to estimate a structural model of parental monetary investments in resources and time on children within a lifecycle model of the household. In their model child quality (human capital) is measured by cognition and parents define their investments in time and resources taking into account the dynamic production function. In my context the human capital vector has three dimensions -health, cognitive skills and socio-emotional skills. However, I do not estimate a model of household decision making. A reason for not doing this is that I do not wish to assume that parents have full knowledge about the production function of human capital [2]. Thus parental decisions are reflected in a reduced form investment equation, of interest in its own right, and the production functions are estimated without imposing that parents know them.

The paper is organized as follows. Section 1.2 presents the conceptual framework of child capability production function. Section 1.3 discusses the empirical analysis of child capability development. In particular, I present four main components of the model: (i) a measurement system to measure child capabilities (subsection (1.3.1)); (ii) a baseline Cobb-Douglas production technology (subsection (1.3.2)); (iii) a translog production technology (subsection (1.3.3)); and (iv) the approach to identify and estimate productivity of

[2] Recent literature shows that subject expectations about the capability production technologies can be altered by policy interventions (Field et al., 2009). Cunha et al. (2013) find that maternal subjective expectations are systematically different from the objective estimates of the production technology. Similarly, Jensen (2010) shows that students underestimate objective returns to schooling and subjective expectations of the returns are extremely lower than the objective measures.

parental monetary investment when it is endogenous (subsection (1.3.4)). Section (1.5) discusses the main estimation results. Section 1.6 presents policy implications and conclusions.

1.2 The conceptual framework of child capability production function

I study the impact of parental monetary investment on the development process of child health, cognitive skills, and socio-emotional skills across multiple stages of childhood. The conceptual framework of capability formation here builds on Heckman et al. (2006), Cunha and Heckman (2007), recently revised in Heckman and Mosso (2014), that I adapt to my context as below.

First, childhood is assumed to consist of multiple development stages, s, with $s = 1, ..., \bar{S}$, which, in principle, includes prenatal time. In my framework, the vector of capabilities consists of three elements, $\Theta_s = (\theta_{H,s}, \theta_{C,s}, \theta_{N,s})$, in which θ_H denotes the stocks of physical health, θ_C denotes cognitive skills, and θ_N for socio-emotional skills[3]. The domain of Θ_s may evolve as the child grows up. Θ_s is a vector of stock variables and is measured at the beginning of each development stage.

In each stage s, parents make investments, I_s, which produce child capabilities in the next stage Θ_{s+1}. The dynamic capability production technology is written as:

$$\begin{aligned} \Theta_{s+1} &= A_s f_s(\Theta_s, I_s) \\ A_s &= a_s(w_s, X_s, \epsilon_{\Theta,s}) \end{aligned} \tag{1.1}$$

where $A_s > 0$ is a dynamic total factor productivity term (TFP) and $f_s(\Theta_s, I_s)$ is a sub-function of the capability production technology. The existing stock of child capability Θ_s and parental monetary investments I_s are inputs to produce the stocks of capabilities of the next stage, $\Theta_{s+1} = (\theta_{H,s+1}, \theta_{C,s+1}, \theta_{N,s+1})$. The TFP term A_s, the technological level at the stage s, depends on family wealth, w_s, other family characteristics X_s, such as caregiver's education and sibling sizes, as well as unobserved shocks $\epsilon_{\Theta,s}$.

I now discuss the implication of Equation (1.1). First, allowing $f_s(.)$ to vary by s implies that the production function may be different at different stages and it accomodates the presence of sensitive and critical periods of child development. With respect to parental monetary investments, the sensitive periods are the times during which investment is most effective in promoting capabilities, $f_s^{'}(I_s) > f_{s'}^{'}(I_{s'})$ with $s \neq s^{'}$. The timing of the sensitive periods is likely different for different capabilities. Second, the term Θ_s as production input captures three ideas: i) the past capability $\theta_{k,s}$ may improve its own stocks at later periods $\theta_{k,s+1}$, termed "self-productivity"; (ii) different elements of the capability vector can (but need not) work synergistically, that is $\theta_{k,s}$ may promote $\theta_{k',s+1}$ with $k \neq k'$ termed as "cross-productivity"; (iii) as long as the components of Θ_s are effective in promoting Θ_{s+1}, parental monetary investment productivity does not fully depreciate within one stage. Third, I allow for the presence of a TFP term which depends on family wealth and family background factors. This captures the possibility that caregivers with higher education are more (less) efficient in promoting child capability through parental monetary investments. The inclusion of family wealth in Equation (1.1) embodies the idea that growing up in poverty does not only undermine parental monetary investment choices and expose children to risk factors but poor parents (facing psychological stress and limited mental bandwidth) may also be less efficient in promoting child development. This also accounts for the fact that poverty can directly impede individual cognitive function and sociobehavioural skills (Mullainathan and Sharif, 2012, 2013; Mullainathan et al., 2013; World Bank, 2018; Haushofer and Fehr, 2014).

[3]socio-emotional skills may consist a very broad range of skills, sometimes loosely called noncognitive skills. Commonly mentioned are grit, self-control, patience, temperament, risk aversion, time preference (Heckman, 2007; Cunha and Heckman, 2007; Heckman and Cunha, 2008, 2010), self-management, effective communication, and prosocial behaviors (Durlak et al., 2011; Heckman et al., 2013, Murnane et al., 2001).

1.3 The empirical analysis of child capability production function

The empirical analysis of production technology I perform largely builds on Cunha et al. (2010), on the recent work by Attanasio et al. (2017a, 2017b, 2018), and on Agostinelli and Wiswall (2016a, 2016b). Specifically, I estimate production technologies in which capabilities are unobserved and measured with error. I also allow for the presence of three important features: (i) Hicks-neutral total factor productivity (TFP) dynamics; (ii) non-constant returns to scale; and (iii) complementarity (or substitutability) between investments and capabilities.

The empirical model consists of five parts. The first is a model of capability development where capabilities in the next stage are produced by the current stocks of child capabilities and parental monetary investment[4]. The second part is a model to approximate parental monetary investment decisions where investments are endogenously determined by the current stocks of child capabilities, family wealth, and family characteristics, and welfare participation. The main difference between the empirical analysis and the conceptual framework in Section 1.2 is that I allow parental monetary investment to endogenously depend on the existing stock of child capabilities, family wealth and characteristics, as well as eligibility for welfare. Moreover, my model entertains the possibility that I know less about the children in the study than they know about themselves. This point will be made explicitly in my analysis of measurement errors in child capabilities and of endogenous investment. The remaining parts of my empirical model are: 3) a model to approximate the initial stocks of child capabilities at the beginning of the development process; 4) a model that relates child capabilities at the final stage of childhood to outcomes at age of majority (expected earnings, expected schooling); and 5) a measurement model to estimate the unobserved latent capabilities from observed measures, allowing for measurement errors.

Because the empirical stages of child development are based on the Young Lives study for Vietnam, I refer to the data features (described in Section 1.4.1) to justify several modelling choices. There are four stages of childhood: age 0-1 ($s = 1$), 2-5 ($s = 2$), 6-8 ($s = 3$), 9-12 ($s = 4$). The stock of existing capabilities is measured at the start of each stage while the flow of parental monetary investment and other inputs are reported for the most recent year of the previous stage. In other word, the coefficients in the production function and investment equations capture averages of input variables on child capabilities from the start to end of the stage as the variables are not measured on an annual basis.

1.3.1 Measuring unobserved capabilities

From Equation (1.1) it is clear that if I could observe the abilities $(\theta_H, \theta_C, \theta_N)$, estimating the production technology is tremendously simplified. Problems arise because I observe only measures of child abilities, which are the external manifestation of skills and so are contaminated by measurement errors[5]. Following the recent approach, I estimate the latent capabilities from a measurement system, taking advantage of the availability of multiple measures for each latent capability in the Young Lives study. The main purpose here is to estimate the latent abilities, not contaminated by measurement errors, and quantify the extent of noise in observed measures. I follow the approach of Cunha and Heckman (2008), Cunha et al. (2010) to estimate the latent capabilities Θ. The basic idea behind this approach is that the analyst can relate observed measures (e.g., cognitive test scores) to unobserved latent capabilities (e.g., cognitive skills) through a measurement system.

Suppose there are $M_{k,s}$ observed measures of child capability $\theta_{k,s}$ at stage s, where $k = \{H, C, N\}$. Specifically, $M_{C,s}$ is the test scores (e.g., reading and math) for cognitive skills, $M_{N,s}$ is the scores of psychological

[4] The Young Lives (YL) data used in my analysis include information about parenting time (of the father and the mother) only in the second wave in 2006. From the third wave in 2009, there is information about how much time a child spent on different activities and whether he/she was supervised by someone. That a majority of the YL families has multi-generations cohabitating makes it infeasible to infer who was the supervisor (parents, grandparents or any other adults living in the same household). Moreover, the dynamics framework of the paper requires that parenting time be available in every stage (three in total). Therefore, I decide not to use this information and focus on parental monetary investment. The same issue incurred in previous work using the YL data, for example, Attanasio et al. (2017, 2018).

[5] See Cunha & Heckman (2008), Kautz et al. (2014) for application in economics. For psychological literature, see for example, Bollen (1989).

scales and $M_{H,s}$ is anthropometric measures and caregiver's assessment about child health. Denote $m_{k,j,s}$ the j-th measure of child's capability k at stage s, $j = 1, ..., J_{k,s}$. I can, therefore, map the j-th measure to the latent capability k as follows:

$$m_{k,j,s} = \alpha_{1,k,j,s} + \alpha_{2,k,j,s} ln\theta_{k,s} + \mu_{k,j,s} \tag{1.2}$$

in which $\alpha_{1,k,j,s}$ is the intercept, $\alpha_{2,k,j,s}$ is the factor loading, and the term $\mu_{k,j,s}$ is measurement error. I assume that each measure $m_{k,j,s}$ is additively separable in the latent factor (in logs) $ln\theta_{k,s}$.In Equation (1.2), $m_{k,j,s}$ is an age-standardized measure of capability k, not a raw score. I consider Equation (1.2) under the assumptions summarized below (Cunha and Heckman, 2008; Cunha et al., 2010; Attanasio et al., 2017a, 2017b, 2018).

Assumption 1. *Measurement model assumptions:*

(a) $\mu_{k,j,s} \perp\!\!\!\perp \theta_{\ell,s'}$ *for all* k, ℓ, *all* j *and all* s *and* s';

(b) $\mu_{k,j,s} \perp\!\!\!\perp \mu_{\ell,j',s}$ *for all* $k \neq \ell$, *all* $j \neq j'$ *and every* s;

(c) $\mu_{k,j,s} \perp\!\!\!\perp \mu_{k,j',s'}$ *for all* $s \neq s'$ *and all* j *and* j';

(d) Dedicated measures: each measure $m_{k,j,s}$ *proxies only one latent factor* $\theta_{k,s}$ *for every* j, *all* k, *and all* s.

Assumption (1a) is that the measurement errors are independent of the latent capabilities. Assumption (1b) implies that the measurement errors are contemporaneously independent across measures and across latent factors. Assumption (1c) means that measurement errors are independent over time. While Assumption (1a)-(1c) are sufficient for nonparametric identification of the latent factors, Cunha et al. (2010) show that Assumption (1b) and (1c) could be relaxed to allow for the presence of correlated measurement errors. Finally,Assumption (1d) requires each measure is the proxy of only one unobserved factor (Gorsuch, 1983, 2003; Attanasio et al., 2017; Cunha and Heckman, 2008; Cunha et al., 2010). Although dedicated measurements are not neccessary for identification, this assumption makes the interpretation of the latent capabilities more transparent.

Moreover, the identification of the measurement system (1.2) also requires some normalizations to set the scale and location of the factors (Anderson and Rubin, 1956) because the latent capabilities have no natural scale and location. I use the following normalizations:

Normalization 1. *Measurement model normalizations:*

(a) The factor mean (in logs) is zero, $E(ln\theta_{k,s}) = 0$, *for all* k *and* s;

(b) The factor loading on one of the measures (say the first measure) of each θ_k *is set to 1, i.e.,* $\alpha_{2,k,1,s} = 1$, *for all* k *and* s.

(c) There exists a common anchoring measure $m_{k,1,s}$ *for each* k *and in all* s.

Normalizations (1a) and (1b) are standard in the literature and used in most previous literature (Cunha and Heckman, 2008; Cunha et al, 2010; Attanasio et al., 2017a, 2017b, 2018). The normalization (1a)-(1b) on Equation (1.2) and the use of age-standardized scores also have important implications for the subsequent estimation of the production technology. Specifically, Agostinelli and Wiswall (2016a, 2016b) show that if the measures $m_{k,j,s}$ are raw scores, e.g., total test scores, BMI, and the analyst imposes the two normalizations above (termed as "re-normalization"), this would implicitly limit the class of estimable technology to log-linear function (Cobb-Douglas). On the other hand, using the age-standardized measures would not imply re-normalization. Normalization (1c) implies that there exists a common anchoring measure, which is repeatedly recorded for each capability in all waves of data collection. The availability of this common measure is crucial in the dynamic settings. The scale of the latent factor is set by the choice of which measurement's factor loading is set to 1, which is salient for the interpretation of the estimates. Agostinelli and Wiswall (2016b) point out that in the dynamic settings, valid inference across time is only possible if each latent factor is scaled in the same way in every period. One way to meet this condition is to normalize each factor on the same measure every period. For example, in this paper, the height-for-age z-score is the common anchoring measure for health, PPVT scores for cognitive skills, and self-esteem scores for socio-emotional skills.

Under the specified assumptions (1a) to (1d), Cunha et al. (2010) show that for each latent factor, i.e., $k \in \{H, C, N\}$ in my context, the distribution of the latent factors, $\{\theta_{k,s}\}_{s=1}^{\bar{S}}$, and of the measurement errors, $\{\mu_{j,k,s}\}_{s=1}^{\overline{S}}$ are nonparametrically identified. Therefore, the joint distribution of $\Theta = \{(\theta_{H,s}, \theta_{C,s}, \theta_{NC,s})\}_{s=1}^{\bar{S}}$ is identified. In the empirical analysis, I maintain that the measurement errors within and between capabilities are independent but do not impose any distributional assumptions on the joint distribution of Θ. As a reminder, the number of measurements, $M_{j,k}$, required to identify the distribution Θ in my empirical exercise is $2L + 1$ with L being the number of latent variables.

These assumptions and normalizations are standard in the current literature. While they appear to be strong, they are actually much weaker and more flexible than the assumptions conventionally imposed in the studies that do not take into account the unobservable nature of individual capabilities. Specifically, this measurement system of capabilities offers a crucial advantage over the alternative approach of using only one single measure and assuming that it is a perfect proxy for each unobserved capability. The latter case amounts to assume $m_{k,j,s} = ln\theta_{k,s}$, which is equivalent to set $\alpha_{1,k,j,s} = 0$, $\alpha_{2,k,j,s} = 1$ and $\mu_{k,j,s} = 0$ for all k, j, s. The measurement system in Equation (1.2), allows for the presence of noisy measures, that is, the measure may capture only a part of the underlying unobserved factor, and therefore, for the possibility that some measures may be more correlated to the latent factors than others, $\alpha_{2,k,j,s} \neq \alpha_{2,k,j',s}$.

1.3.2 The baseline technology with TFP dynamics and nonconstant return to scale

After the joint distribution of latent factors $\Theta = \{(\theta_{H,s}, \theta_{C,s}, \theta_{NC,s})\}_{s=1}^{\bar{S}}$has been identified, I proceed to specify an empirical model for the joint development of health, cognitive skills and socio-emotional skills using the following Cobb-Douglas production technology, with addition of TFP dynamics and allowing for nonconstant returns to scale:

$$\begin{aligned} \theta_{k,s+1} &= A_{k,s}\left(\theta_{H,s}^{\gamma_{k,H,s}}\theta_{C,s}^{\gamma_{k,C,s}}\theta_{N,s}^{\gamma_{k,N,s}}I_s^{\gamma_{k,I,s}}\right) && \text{with k} \in \{H,C,N\}\text{,s=1,2,3,4} \\ A_{k,s} &= exp\left(\gamma_{k,0,s} + \gamma_{k,w,s}w_s + \gamma_{k,X,s}X_s + \epsilon_{k,s}\right) && \text{with } A_{k.s} > 0 \end{aligned} \tag{1.3}$$

in which, $A_{k,s}$ is the TFP terms, which depends on family wealth w_s and family background factors X_s. Rewriting in a log linear fashion and rearranging, Equation(1.3) becomes:

$$\begin{aligned} ln\theta_{k,s+1} &= \gamma_{k,0,s} + \textstyle\sum_{\ell}\gamma_{k,\ell,s}ln\theta_{k,s} + \gamma_{k,I,s}lnI_s + \gamma_{k,X,s}X_s + \epsilon_{k,s} \\ & \qquad \text{with } k, \ell \in \{H, C, N\} \end{aligned} \tag{1.4}$$

in which $\epsilon_{k,s}$ denote stage-specific unobserved shocks to the development process. It is important to emphasize that the specification does not ignore the genetic or very early contribution (prenatal) of parents, but does assume that they are captured in the initial conditions at stage $s = 1$ (age 0-1) - a first order Markov-type assumption.

I impose the following assumptions on the technology in Equation (1.4) (Cunha and Heckman, 2008; Cunha et al., 2010; Attanasio et al., 2017a, 2017b, 2018; Agostinelli and Wiswall, 2016a, 2016b):

Assumption 2. *Production technology assumptions:*

$\epsilon_{k,s} \perp\!\!\!\perp \mu_{k,j,s}$ *for all* k*, all* j*, and all* s*.*

Assumption (2) requires that the idiosyncratic shocks to the development process of child capabilities are independent of the errors in measuring those capabilities. This is a standard assumption imposed in all previous work in the literature that acknowledges the unobserved nature of child capabilities.

My baseline analysis of the technology of capability production function departs from previous work in several ways. (1) Previous work using the multistage framework have focused on modelling the joint evolution of cognitive skills, and socio-emotional skills (Cunha and Heckman, 2008; Cunha et al., 2010; Helmers and Patnam, 2011), of health and cognitive skills (Attanasio et al., 2017a, 2017b), or of cognitive skills alone (Todd and Wolpin, 2005; Del Boca et al., 2014; Agostinelli and Wiswall, 2016a, 2016b). As to my best knowledge, this paper is the first attempt to model health, cognitive skills, and socio-emotional skills simultaneously over the childhood.

(2) I allow the production technology in Equation (1.4) to exhibit Hicks neutral TFP dynamics, i.e., $A_{k,s} > 0$, and nonconstant returns to scale (non-CRTS), that is $\sum_{\ell} \gamma_{k,\ell,s} + \gamma_{k,I,s} \neq 1$ with $k, \ell \in \{H, C, N\}$. Among previous work, only Agostinelli and Wiswall (2016a, 2016b) have considered a class of CES technology which exhibits nonconstant return to scale. They point out that allowing for TFP dynamics and non-CRTS is empirically important and can substantially change the inferences about the child development process. However, the two studies only investigate the development of cognitive skills. In my context of the multi-dimensional capability vector, I find that TFP dynamics play a significant role in the development of child capabilities and that there is evidence of decreasing returns to scale for all capability production functions during childhood.

(3) As discussed in detail in Section 1.3.4.3, I account for the endogeneity of parental monetary investments, which result from parental responses in investments to development shocks to child capabilities observed by parents but unobserved to the analyst. That is, I allow the stochastic production shock, $\epsilon_{k,s}$ to be correlated with the flow of monetary investments, $I_{s.}$ Accounting for investment endogeneity substantially affects the estimates of the production technologies. To my best knowledge, this is the first study that allows for the aforementioned features (1) and (2) and simultaneously uses instrumental variables to correct for investment endogeneity.

I highlight hereafter some implications of the production technology in Equation (1.4). First, the stage-specific productivity of parental monetary investment on capability θ_k ,$\gamma_{k,,s}$, $k \in \{H, C, NC\}$, captures the presence of sensitive periods. Specifically, the periods of high plasticity, or sensitive periods, whose length and presence may differ widely by capability, are characterised by a higher level of malleability of capabilities. Given the focus on parental monetary investments in my context, sensitive periods are defined as the periods in which investment is most effective in producing capabilities. The timing of sensitive periods for investments is an empirical question which I investigate in this analysis.

Second, as discussed above, the inclusion of current capabilities $\theta_{k,s}$, $k \in \{H, C, N\}$, captures two ideas: (i) different elements of the human capital vector can be cross-productive in promoting future stocks of capabilities, i.e.,

$$\frac{\partial \theta_{k,s+1}}{\partial \theta_{\ell,s}} > 0 \;\; \text{with } k \neq \ell \text{ and } \ell \in \{\text{H,C,N}\};$$

(ii) the effects of past investment and past time-variant family characteristics on child development may not fully depreciate after one stage, i.e.,

$$\frac{\partial \theta_{k,s+1}}{\partial I_{s'}} > 0, \quad \text{with } s^{'} \leq s.$$

Regarding the first implication, the epidemiology, neuroscience and economics literature has shown that poor health (stunting and infectious diseases) may inhibit or hinder normal brain development, and thus, early cognitive development (Knudsen, 2004). Similarly, emotional security promotes exploration, which paves the way to cognitive development; and better self-control lowers the incidence of unhealthy behaviours (Heckman, 2007; Knudsen et al., 2006). While the cross-productivity among health, cognitive skills, and socio-emotional skills are widely documented; it is not clear whether their coevolution results either from the exposure to common risk factors (or stimuli) or from pure cross-productivity or both. The distinction between them is important because different causes imply different policy intervention approaches. For example, if the observed health and socio-emotional cross-productivity on cognitive development in disadvantaged children results from the exposure to common risk factors, interventions that improve living conditions or investments (parental or societal) alone can fully boost child cognition. Otherwise, an intervention without health- and socio-emotional-related elements would not fully realize child potentials.

1.3.3 The translog technology with nonconstant eslasticity of subsitution

As discussed in Section (1.3.2), the technology (1.4) does not allow for the possibility of nonzero, nonconstant elasticity of substitution among inputs (Θ_s, I_s). Therefore, in addition to Equation (1.4), I estimate a stochastic translog production function as follows:

$$\begin{aligned} ln\theta_{k,s+1} &= lnA_{k,s} + \sum_\ell \gamma_{k,\ell,s} ln\theta_{\ell,s} + \gamma_{k,I,s} lnI_s + \sum_\ell \beta_{k,\ell,s} lnI_s ln\theta_{\ell,s} \\ lnA_{k,s} &= \gamma_{k,0,s} + \gamma_{k,w,s} w_s + \gamma_{k,X,s} X_s + \epsilon_{k,s}. \end{aligned} \tag{1.5}$$

The translog technology can also be expanded to include additional terms which are capable of approximating many unknown production technologies. As before, $A_{k,s}$ is the TFP term and $\epsilon_{k,s}$ is the idiosyncratic production shock assumed to be i.i.d with $\epsilon_{k,s} \sim N(0, \sigma^2_{k,s})$ for all s and independent of all inputs (Θ_s, I_s). The term $\epsilon_{k,s}$is also assumed to be independent of the measurement errors in Equation (1.2).

The translog technology in Equation (1.5) allows for nonconstant returns to scale and for nonconstant elasticities of substitution among inputs (Θ_s, I_s). Therefore, it enables me to investigate whether child capabilities and investment are complements or substitutes. With $\beta_{k,\ell,s} = 0$ for all k, ℓ, s, the specification (1.5) reduces to the specification (1.4). With $\beta_{k,\ell,s} \neq 0$, the elasticity of capability production with respect to parental monetary investment is dependent on the current stock of child capabilities:

$$\frac{\partial ln\theta_{k,s+1}}{\partial lnI_s} = \gamma_{k,I,s} + \sum_\ell \beta_{k,\ell,s} ln\theta_{\ell,s} \quad \text{with } k, \ell \in \{H, C, N\},$$

in which $\beta_{k,\ell,s} > 0$ would imply a higher return to parental monetary investment on capability k for children with higher current stocks of capability ℓ than for those with lower current stocks of capability ℓ, and vice versa if $\beta_{k,\ell,s} < 0$. Moreover, as long as $\beta_{k,\ell,s} \neq 0$, the productivity of I_s in producing $\theta_{k,s+1}$ will also depend on the past stocks of capability (and thus, past parental monetary investments) at stages $s' < s$. A similar argument also applies to the elasticity with respect to current stocks of capabilities:

$$\frac{\partial ln\theta_{k,s+1}}{\partial ln\theta_{\ell,s}} = \gamma_{k,\ell,s} + \beta_{k,\ell,s} lnI_s.$$

For example, $\beta_{k,\ell,s} > 0$ suggests that self-productivity of child capability would be higher for children with higher levels of current capability and receiving more investments.

1.3.4 Using welfare eligibility to estimate the effects of investment

Until now I have assumed that the unobserved shocks in the production function of capability k, $\epsilon_{k,s}$, are independent from parental monetary investments, I_{s-1}, as well as from the unobserved shocks in the technology of capability $k' \neq k$. In other words, the non-monetary production inputs, which are not observed by the analyst, do not have any impact on parental monetary investments. In practice, parents may change their monetary investments in response to the inputs in $\epsilon_{k,s}$ which are not observed by the analyst. I now discuss an approach for identifying the investment coefficients in Equation (1.4) with endogenous parental monetary investment, exploiting the eligibility to participate in welfare programs as a source of exogenous variation.

1.3.4.1 Housing support programme for the poor in Vietnam

I use eligibility for welfare participation as instrumental variables to help identify the model.The welfare rules I consider allow me to take advantage of a national housing support programme (HSP) during the 2000s and 2010s that entitled *every* poor households living in rural areas to unconditional cash transfers for improving living conditions and/or low-interest loans at the national scale. This housing programme is under the umbrella of National Programmes for Poverty Reduction, to which the impressive poverty reduction rates in Vietnam, from 39.9% (1993) to 4.1% (2008)[6], are often attributed. A crucial advantage offered by these instruments is that they induce variation in the parental monetary investments across space and over time for a single cohort of children, allowing for a very tight identification strategy to handle the endogeneity of investment.

[6]Using the World Bank poverty line of $1/day.

The HSP consists of multiple components with different demographic and geographical targets, different levels of benefits (i.e., how much cash transfers and loan eligible families receive), as well as timing of implementing and revising schedules. Specifically, the HSP has been part of three different poverty alleviation programmes implemented since 1998 to 2020 (potentially): (i) the housing support component of the National Targeting Program for Poverty Alleviation (NTP) from 1998-2020[7]; (ii) the support for access to land, housing and access to water programme (P134) implemented from 2004 with the Decision 134/2004 and Decision 167/2004; (iii) the reallocation component of the Program 135 for Ethnic minorities from 1998-2010. While these policies have different geographical coverage, they share a common ground for core eligibility criteria, which are deprived living conditions. They also share common benefit rules as follows. HSP beneficiaries receive an amount of unconditional cash transfer and/or low interest loan which are subject to change across provinces and over time. Once HSP households receive cash transfers, housing construction must take place soon and local authorities monitor and ensure that new houses meet standards set by the provincial authority. These households are also automatically ineligible for any HSP benefits in the future. This feature of the HSP programme creates further innovation in household eligibility over time, strengthening the identification strategy of my model.

1.3.4.2 Eligibility for the Housing Support Programme

In principle, the actual benefits (HSP participation) shift the household budget upwards due to reception of the unconditional cash transfers, which must be spent on improving housing conditions. Consequently, the HSP benefits may induce two distinct effects: (i) it may relax liquidity constraints for the constrained families; and (ii) by improving the housing conditions, it may have a direct impact on child capabilities, especially child health. The second situation would invalidate the use of actual HSP participation as an instrument for parental monetary investments.

To circumvent these difficulties when both actual HSP participation and actual eligibility are prone to manipulation, I seek to simulate the HSP eligibility scores. I expect that constrained parents, who are unable to insure against their own income and likely make suboptimal investments (conditional on children's existing capabilities) to change investment decisions under the prospect of being eligible for HSP benefits. Moreover, by using simulated eligibility instead of actual eligibility, I could safely sidestep the possibility that the latter could be either the occurence of propitious selection or result from families' manipulation or resource misallocation by local authorities.

Specifically, I compute simulated HSP scores for each family over three periods, 2001-2006, 2006-2009, and 2009-2013, following the official guidelines of the HSP programme. To fix the idea, I reintroduce the subscript i to denote child i from family i and write the HSP eligibility variable, $HSP_{i,s}$, as a function of housing conditions at stage s, which are essentially parts of the family wealth index, $w_{i,s}$, and other permanent characteristics, including ethnicity and living location (rural/urban). The HSP calculation is given by:

$$HSP_{i,s} = \chi_s(w_{i,s}, X_i) = \sum_{l=1}^{6} \mathbf{i}_\ell(c_{i,s,l} \leq \bar{c}_s) \text{ with } \mathbf{i}_l = \begin{cases} 1 & \text{if } c_{i,s,l} \leq \bar{c}_{s,l} \\ 0 & \text{otherwise} \end{cases} . \tag{1.6}$$

Specifically, $HSP_{i,s}$ is the sum of six indicator variables, $\mathbf{i}_l$, $l = \{1, .., 6\}$, corresponding to six selection criteria for actual HSP eligibility, which I collected from official guidelines of the program. The term $\mathbf{i}_l$ takes value of one if a physical housing condition does not meet the threshold, i.e, $c_{i,s,l} \leq \bar{c}_{s,l}$, and zero otherwise. The maximum score, $HSP_{i,s} = 6$, implies that family i is most likely to benefit from the HSP.

Given the construction of the simulated HSP eligibility, I now turn to the discussion of its validity as instrumental variable. The first concern over its validity is whether the simulated HSP eligibility is correlated with unobserved inputs of the production function. For example, while this study focuses on material investment,

[7]The first phase of the NTP was from 1998 to 2000 under the Decision 133/1998; the second phase ran from 2001 to 2005 under the Decision 143/2001; the third phase from 2006 to 2010 and implemented following the Decision 20/2005. In 2012, the NTP changed fundamentally its overarching target from poverty alleviation to sustainable poverty alleviation for the period of 2012-2020 following the Resolution 80/NQ-CP (2011) and the Decision 1489 (2012). The fourth phase in 2012-2015 was made effective under Decision 1489 (2012). The fifth phase starts from 2016-2020 following the Decision 1722 and uses multidimensional measurement of poverty.

I cannot rule out *ex ante* the possibility that parental time investment, which is omitted in the empirical exercise, may be subject to changes in response to their HSP eligibility status. I present evidencewhich rules out this possibility in the robustness checks. Secondly, I must also rule out that conditioning on $(\Theta_s, I_s, w_s, X_s)$, being eligible for the HSP benefits at stage s has direct effects on child capabilities. As I already discussed, the wealth effects of HSP eligibility, or the prospect of budget constraint shifted upward, is unlikely to be a direct one on child development, apart from potential improvement in living conditions. Morover, the set of variables that is used to construct HSP eligibility is a subset of indicators used to construct the wealth index. Therefore, any actual changes in wealth if ever made prior to the actual reception of HSP benefits should be fully captured in the estimates of wealth coefficients.

1.3.4.3 Correcting for investment endogeneity in the Cobb-Douglas production function

This section discusses the correction for investment endogeneity in the Cobb-Douglas production function in Equation (1.4). I follow the approach by Cunha et al. (2010), Attanasio et al. (2017a, 2017b), Agostinelli and Wiswall (2018), and Bernal and Keane (2011) to estimate a tractable approximation of the parental monetary investment decision. The idea is to approximate parental monetary investment as a function of the child capabilities, family background factors, and to use HSP eligibility as stage-specific instrumental variables.

First, I assume parental monetary investments depend on: (i) current stocks of child capabilities; (ii) caregiver's characteristics; (iii) family wealth; (iv) eligibility for welfare programs; (v) other child characteristics; and (vi) a shock $\epsilon_{I,s}$ that is unobserved to the analyst. A conditional decision rule for monetary investment I_s is given by:

$$I_s = g\left(\Theta_s, w_s, HSP_s(w_s, X_s; R), X_s, \epsilon_{I,s}\right)$$

in which welfare rules R govern how potential HSP benefits depend on stage-specific family wealth w_s and family background X_s, and HSP_s is family simulated eligibility to participarte in the HSP programme. Endogeneity of investment implies non-zero correlation of the investment error components $\epsilon_{I,s}$ with each of the $\epsilon_{k,s}$ for all $k \in \{H, C, N\}$.

The decision of parental monetary investments is empirically approximated by the function below:

$$lnI_s = \quad \gamma_I + \textstyle\sum_k \gamma_{I,k} ln\theta_{k,s} + \gamma_{I,X} X_s + \gamma_{I,w} w_s + \gamma_{I,HSP} HSP_s + \epsilon_{I,s} \tag{1.7}$$

with $E(\epsilon_{I,s}|HSP_s) = E(\epsilon_{k,s}|HSP_s) = 0$. The identification of investment coefficients in the production technology (1.4) is then straightforward, using standard arguments for instrumental variables with HSP_s serving as stage-specific instruments for I_s in each stage s. This specification should be seen as a "conditional investment decision rule" of the more complex analytic decision rule obtained from a dynamic optimization process (Bernal & Keane, 2011; Attanasio et al., 2015a)[8]. Because I do not derive the investment decisions from an explicit economic model, the interpretation of estimated coefficients in Equation (1.7) is rather speculative.

Because HSP_s reflects the extent to which the child's family is entitled to the HSP benefits, I expect that conditional on family wealth, w_s, and family background, X_s, the higher values of HSP_s would induce parents to invest more in their children, i.e., $\gamma_{I,HSP} > 0$. It is also important to emphasize that my analysis deviates from previous studies, using the same framework of multi-dimensional child capabilities and multi-stages of development, in that I allow family financial resources (wealth) to be a direct input of the development process, not just indirectly through their effects on parental monetary investments[9]. Therefore, contrary to

[8] There are two prespective on this feature of the model: (a) as an incomplete specification because it permits a number of behaviours of parents though the original economic model does not; or (b) a desired property because it allows for richer (and more general) economic behaviours (Heckman and Navarro, 2007).

[9] For example, a range of problems from physical and mental deficiency to cognitive impairment are direct consequences of the early experience of growing up in noxious, stressful environments (captured in my wealth indices) (Evans and Kim, 2013; McCoy and Raver, 2014).

the other studies, I do not consider family wealth as a possible instrumental variable for endogenous parental monetary investments[10].

Turn to the interpretation of the coefficients in Equation (1.7). First of all, it is crucial to emphasize that the coefficients are likely to capture combined effects of covariates on parental monetary investments originated by several mechanisms. For example, the response of parents to potential stimulus/benefits - welfare benefits, might reinforce ($\gamma_{I,HSP} > 0$) or compensate ($\gamma_{I,HSP} < 0$) their own monetary investments. Moreover, $\gamma_{I,HSP}$ must also capture average reponses of parental monetary investments to welfare eligibility over the childhood. The term $\gamma_{I,w}$ reflects the effect of family financial resources on the extensive margin of parental monetary investments, which might be combined effects of liquidity constraints[11] and parental preferences in making investments on child capabilities (for example., in deprived conditions even purchased food are considered an investment while in wealthy families, it is consumption (Caucutt et al., 2015)).

1.3.4.4 Correcting for investment endogeneity in the translog production function

Cunha et al. (2010) has demonstrated that given a stage-specific instrument (in my case HSP_s), a general class of capability production functions can be identified, including the translog function as in this paper (see Cunha et al., 2010, for details). They consider family resources as instrumental variables for endogenous parental investments and to identify the technology parameters, using a control function approach to estimate for their empirical analysis. The recent studies by Attanasio et al. (2017a, 2017b) also follow the control function approach. While their identification analysis directly applies to the case of translog production function, my empirical analysis follows a different approach to estimate the production function, which is discussed in detail in Woolridge (2010) with original arguments traced back to Fisher (1965).

I start by rewriting the production and investment system. Without loss of generality, I assume that the human capital vector is of single dimension, and write the translog specification in Equation (1.5) as:

$$ln\theta_{s+1} = \gamma_{1,s} ln\theta_s + \gamma_{2,s} lnI_s + \gamma_{3,s} ln\theta_s lnI_s + \epsilon_{\theta,s}. \tag{1.8}$$

in which I define the interaction term $z_s \equiv ln\theta_s lnI_s$. Given that investment I_s is likely to be correlated with $\epsilon_{\theta,s}$, the variable $\boldsymbol{z_s}$ is also endogenous. The approximate investment decision is also rewritten as:

$$lnI_s = \gamma_{I,1} ln\theta_s + \gamma_{I,2} HSP_s + \epsilon_{I,s} \tag{1.9}$$

with $E(\epsilon_{I,s}|HSP_s) = E(\epsilon_{\theta,s}|HSP_s) = 0$. For simplicity, the TFP terms in the production function, the intercepts in the investment equation, and all other covariates are not shown explicitly. The eligibility for HSP, HSP_s, shifts the investment decision but not the capability production technology. For each stage $s+1$, Equation (1.8) and (1.9) constitutes a nonlinear system in endogenous parental monetary investment. However, they are still linear in parameters, which enables us to estimate the whole system using the common instrumental variable arguments.

There are multiple possible instruments available for $lnI_s ln\theta_s$. Because $E(lnI_s ln\theta_s|HSP_s)$ is not linear in HSP_s even if $\gamma_{3,s} = 0$ in Equation (1.8), other functions of HSP_s will be the independent variables in a linear projection with $\boldsymbol{z}_s$ as the dependent variable. To see these instruments, first, suppose that $\gamma_{3,s} = 0$, which is the Cobb-Douglas function I already discussed above. Multiplying both sides of Equation (1.9) gives

$$E(\boldsymbol{z}_s|HSP_s) = \gamma_{I,1} ln\theta_s ln\theta_s + \gamma_{I,2} HSP_s ln\theta_s + E(\epsilon_{I,s}|HSP_s)E(ln\theta_s|HSP_s) \tag{1.10}$$

where the last terms is dropped out using $E(\epsilon_{I,s}|HSP_s) = 0$. Equation (1.10) shows that $\boldsymbol{z}_s$ is correlated with $HSP_s ln\theta_s$, making it a natural instrument for $\boldsymbol{z}_s$ in Equation (1.8). The only case when $HSP_s ln\theta_s$ is

[10] Glewwe et al. (2017) investigate the production function of cognitive skills in Vietnam, using the Young Livesdataset. They find that after controlling for the educational inputs, and child, parental and family background, the correlation between wealth and child cognitive test scores is not significant. One possible explanation for the result is that the authors use test scores without correcting for associated measurement errors which in turn will bias the wealth coefficient.

[11] The liquidity constraints can be further decomposed into three types: (i) the liquidity constraint on parents to insure against their child's future income, (ii) the liquidity constraint on children to insure against their own future income, (iii) the "accident of birth" which means children cannot self-select into a particular family (Cunha et al., 2006; Cunha and Heckman, 2007; Heckman and Mosso, 2014).

not correlated with z_s occurs when $\gamma_{I,HSP} = 0$, in this case, the Cobb-Douglas production function (with $\gamma_{3,s} = 0$ in Equation (1.8)) is also unidentified.

In practice, I augment Equation (1.8) with the following linear projections:

$$\begin{array}{rcl} lnI_s & = & \gamma_{I,1}ln\theta_s + \gamma_{I,2}HSP_s ln\theta_s + \gamma_{I,3}HSP_s + \epsilon_{I,s} \\ \boldsymbol{z}_s & = & \gamma_{z,1}ln\theta_s + \gamma_{z,2}HSP_s ln\theta_s + \gamma_{z,3}HSP_s + \epsilon_{z,s} \end{array} . \tag{1.11}$$

This also suggests a sequential approach in the empirical analysis: (i) estimating the baseline Cobb-Douglas technology jointly with the investment equation (1.7) and using HSP_s as an instrument variable for I_s; (ii) testing the null hypothesis that $\gamma_{I,HSP} \neq 0$; and (iii) estimating the translog technology jointly with the system (1.11) and using HSP_s and $HSP_s ln\theta_s$ as instruments for I_s and z_s, respectively.

1.3.5 Anchoring child capabilities and initial conditions

1.3.5.1 Anchoring of child capabilities

It is established in the literature of child development that cognitive test scores, psychological measures, and health measures, such as maternal assessment of child health, are arbitrarily scaled. The problem is due to the fact that any monotonic transformation of the measures are also conceptually valid measures of child capabilities (Cunha et al., 2010; Bond and Lang, 2013, 2017; Nielsen et al., 2015). This is resolved neither by standardizing test scores nor by applying more complex methods to calculate the scores, such as the item response model. Therefore, an important issue related to the scale of the latent factors is how they actually contribute to later-life outcomes, suh as wages or unhealthy behaviors. To investigate the relative importance of child health, cognitive skills, and socio-emotional skills and of parental monetary investments over different periods of childhood, I augment the production technology and investment equations with an equation for child's expected schooling at majority age, following Cunha and Heckman (2008) and Cunha et al. (2010). The expected schooling variable provides a common, interpretable scale for arbitrarily scaled latent cognitive skills and socio-behavioural skills. The outcome Q is measured without errors and given by:

$$Q = \gamma_{H,Q} ln\theta_{H,S} + \gamma_{C,Q} ln\theta_{C,S} + \gamma_{NC,Q} ln\theta_{NC,S} + \epsilon_Q \tag{1.12}$$

where S is the last development stage. In the case of endogenous parental monetary investments, I allow ϵ_Q to be correlated with investment errors at $s = S$. The outcomes Q are measured at age 16 - the official age of majority in Vietnam. Expected schooling has been long established as among the most important determinants of college enrolment (Card, 2001), which, in turns, raises both individual realized earnings and other nonpecuniary outcomes in later life.

1.3.5.2 Initial conditions of the development process

Finally, to complete the model, I specify an equation for the initial stocks of child capabilities, Θ_1. Ideally, I should have measures of each capability endowment, $\Theta_1 = (\theta_{H,1}, \theta_{C,1}, \theta_{NC,1}),\ k \in \{H, C, NC\}$. However, the Young Lives study does not include specific measures for child cognitive before age 1 and before age 8 for sociobehavioural skills. Therefore, I assume that the child's initial endowment at $s = 1$ can be summarized into a single latent variable Θ_1 and given by:

$$ln\Theta_1 = \beta_1 \mathbf{Z}_1 + \epsilon_1, \tag{1.13}$$

in which $\boldsymbol{Z}_1$ is the vector of family-level variables, Θ_1 is also a latent capability and estimated according to the measurement procedure discussed in Section 1.3.1. The initial capability is correlated with (i) the maternal observed characteristics - her age at the time of childbirth and ethnicity; (ii) primary caregiver's education; and (iii) family characteristics such as family wealth and sibling sizes. Though the monetary investment variable does not enter Equation (1.13), by incorporating these characteristics of the family and the child, I do not strictly assume that prior to the beginning of the sample period, investment activities are nonexistent. Finally, note that Equation (1.13) does not necessarily captures any causal relationship between $\boldsymbol{Z}_1$ and Θ_1, but may merely reflect correlation. The main purpose is to let $\boldsymbol{Z}_1$ "absorb" as much of the child unobserved initial capability, ϵ_1, as possible, because this would lower the sensitivity of the results to the assumptions on the whole system.

1.3.6 Estimation procedure

All the necessary information to analyze the production technology is embodied in the joint distribution of the latent factors, parental monetary investments, simulated HSP eligibility and key family background factors. My approach to estimation[12] of the technology parameters consists of two steps, closely following the three steps of identification above.

The first step uses a minimum distance estimator to estimate the distribution of latent skills, denoted $\mathcal{F}(\Theta)$, and to project the full sequence of latent variables over time. The estimator[13] is robust and does not impose any distributional assumptions on the distribution of latent skills and measurement errors. This flexibility in the measurement systems is especially crucial for estimating the translogproduction technologies. Specifically, the presence of interactions between Θ_s and I_s in the translog technology implies nonlinear conditional means of the latent factors , which, in turns, precludes the assumption of multivariate normality of Θ_s.

In the below, I summarize the estimation steps of the measurement model (1.2) using weighted least square estimator. This WLS estimator is robust, tractable and widely used in the psychometric literature (Browne, 1982, 1984; Bollen, 1989). First, I rewrite the measurement system (1.2) in the vector form as:

$$\boldsymbol{m} = \boldsymbol{\alpha} ln\Theta + \mu \tag{1.14}$$

Let $\boldsymbol{\eta} = (\boldsymbol{\alpha}, \boldsymbol{\mu})$ denote the vector of parameters of the measurement model. Let $\varSigma_{mm}$ denote covariance matrix of measures $\boldsymbol{m}$ and $\varSigma_{mm}(\boldsymbol{\eta})$ contain the covariance of $\boldsymbol{m}$ as a function of the unknown parameters $\boldsymbol{\eta}$. $\boldsymbol{\Sigma}_{\mu}$ is the covariance matrix ofthe measurement errors collected in μ. I can express the covariance matrix of $\boldsymbol{m}$ as a function of of unknown parameters $\boldsymbol{\eta}$, that is:

$$\begin{aligned} \Sigma_{mm}(\boldsymbol{\eta}) = E(\boldsymbol{m}\boldsymbol{m}^{'}) \quad &= E\left[(\boldsymbol{\alpha} ln\Theta + \mu)\left(\boldsymbol{\alpha}^{'}(ln\Theta)^{'} + \mu^{'}\right)\right] \\ &= \boldsymbol{\alpha} E(ln\Theta(ln\Theta)')\boldsymbol{\alpha}^{'} + \varSigma_{\mu} \end{aligned} . \tag{1.15}$$

The WLS fitting function is:

$$F_{MDE} = [\boldsymbol{s} - \sigma_{mm}(\boldsymbol{\eta})]^{'} \boldsymbol{W^{-1}} [\boldsymbol{s} - \sigma_{mm}(\boldsymbol{\eta})], \tag{1.16}$$

in which $\boldsymbol{s}$ is the sample analog of $\sigma_{mm}(\boldsymbol{\eta})$, $\boldsymbol{W^{-1}}$ is a positive-definite weight matrix. The estimates of unknown parameters $\boldsymbol{\eta}$ minimize the weighted sum of squared deviation of $\boldsymbol{s}$ from the $\sigma_{mm}(\boldsymbol{\eta})$. The estimated $\hat{\boldsymbol{\eta}}$ is consistent when $\varSigma_{mm} = \varSigma_{mm}(\boldsymbol{\eta})$ under very general conditions (Browne, 1982, 1984). If the weight matrix $\boldsymbol{W}$ is chosen to be equal to or to be a consistent estimator of the asymptotic covariance matrix of $\boldsymbol{s}$, then the estimated parameters $\hat{\boldsymbol{\eta}}$ from F_{WLS} fitting function (1.16) is asymptotically efficient. Bollen (1989) shows that the optimal choice for the weight matrix $\boldsymbol{W}$ is the covariance matrix of the sample covariances.

Having estimated the parameters $\hat{\boldsymbol{\eta}}$ of the measurement system, I project the whole sequence of health, cognitive skills and socio-emotional skills over $\overline{S}$ stages of childhood ($\{\theta_{H,s}\}_{s=1}^{\bar{S}}, \{\theta_{C,s}\}_{s=2}^{\bar{S}}, \{\theta_{N,s}\}_{s=3}^{\bar{S}}$). The projected latent skills $\hat{\Theta}$ are subsequently used in the estimation of the production technology. Note that the Young Lives study does not include information about child cognitive skills at the first stage (age $0-1$) and about socio-emotional skills at the first and second stage (age $0-1$ and age $1-5$), so the subsequence of cognitive skills starts at $s=2$ and at $s=3$ for socio-emotional skills.

In the second step, given the full sequence of the latent factors, I proceed to estimate the production technology in one of three following cases: (i) the Cobb-Douglas technology (1.4) with exogenous parental monetary investment; (ii) the Cobb-Douglas technology (1.4) with endogenous parental monetary investment which is due to unobserved heterogeneity; and (iii) the translog technology with parental monetary investment endogenously depend on child capabilities, parental backgrounds, and simulated HSP eligibility, and unobserved heterogeneity. In all of the three cases, I estimate simultaneously the following equations: the initial condition (1.13), the production technology (1.4) (or alternatively (1.5)), the anchoring outcome (1.12). When investment is endogenous, I augument the system by jointly estimate these equations with the investment Equation (1.7) (Cobb-Douglas technology), or with Equation (1.9) and (1.10). I allow the error components of the investment equations to be correlated with the error components in the production technology of each latent capability.

[12] I estimate the model using the SEM command in STATA version 14.2 (StataCorp, 2015).
[13] discussed at details in Bollen (1989) and Browne (1982, 1984)

Table 1.1: Data availability

Variable	Description	Study years	Model stages (s)
$\{\theta_{H,s}\}$	health measures	2002, 2006, 2009, 2013	1,2,3,4
$\{\theta_{C,s}\}$	cognitive measures	2006, 2009, 2013	2,3,4
$\{\theta_{N,s}\}$	socio-emotional skills measures	2009, 2013	3,4
I_s	monetary investment	2006, 2009, 2013	2,3,4
W_s	wealth index		
	(housing physical environment)	2002, 2006, 2009, 2013	1,2,3,4
X_s	demographic characteristics	2002, 2006, 2009, 2013	1,2,3,4
$\{HSP_s\}$	simulated eligibility for HSP	2006, 2009, 2013	2,3,4

1.4 Data and descriptive statistics

1.4.1 The Young Lives study

I use the sample of young cohort from the Young Lives (YL) study for Vietnam from 2002 to 2013. The study tracks around 2000 children and provides repeated measures on child's health, cognitive skills and socio-emotional skills over time. These YL children were about 0-1 year old at the first wave in 2001. Follow-up studies have been conducted in 2006, 2009 and 2013.

The YL data are ideal for studying the impact of parental monetary investment on the development of child capabilities. First and most importantly, multiple measures of each dimension of child capability are available from 2002 to 2013 on four occasions, allowing me to identify and estimate the unobserved latent capabilities, not contaminated by measurement errors. Specifically, a health assessment is conducted in all four study waves, providing anthropometric measures and a caregiver's assessment on child health. The health information is highly comparable and similar across waves. Anthropometric z-scores are age-in-days adjusted and normalized to have a mean of zero and standard deviation of one based on a random sample of Vietnamese children. Cognitive tests were administered starting from the second wave in 2006, and psychological questions for socio-emotional skills starting in 2009. I use the scores from the IRT models with mean zero and a standard deviation of one based on the full YL sample of the Vietnamese younger cohort.

Second, the measure of parental monetary investment in this study is child-specific, that is, parents provide information on how much they spend on each YL child on a yearly basis (for the previous twelve months) starting from 2006. The measure of parental monetary investment consists of spending on three main categories: education-, health- and entertainment-related[14]. The total sum of monetary values is deflated to 2006 Vietnam dong and adjusted for rural-urban differences. Third, apart from child-specific, the YL data provide information on family wealth[15] for every wave and these data are used to construct the wealth variable. The inclusion of family wealth, representing family long-term resource availability as well as housing conditions, is crucial for my analysis. Moreover, I use the subsets of information on housing conditions and access to services to construct the simulated eligibility to participate in HSP.

I restrict the main sample to children who were observed in all four study waves and who hold nonmissing values for all capabilities measures, monetary investment, family background factors and HSP simulated eligibility. After dropping missing cases, my main sample includes 1142 children residing in 1142 families. Data used in the empirical estimation are summarized in Table 1.1.

[14] Parents provide information on how much they have spent on YL child on education and healthcare. For other categories, such as gifts, clothing, they are asked about the total spending on all children in the family and the corresponding shares of YL child.

[15] The wealth index is a continuous measure varying within the unit interval of $[0, 1]$ andconsists of three equally-weighted sub-indexes: (i) housing quality, (ii) access to services and (iii) consumer durables (Obiageri & Kristine, 2016).

Table 1.2: Descriptive statistics of parental monetary investment

Model stage	study year	Median child spending ('000 VND)				Median investment	
		total	education	health care	entertainment	('000 VND)	($ equiv.)
2 ($2 \rightarrow 5$)	2006	754.398	245.515	141.643	94.429	2924.739	190.108
3 ($6 \rightarrow 8$)	2009	1162.294	501.812	196.999	16.417	3436.648	223.382
4 ($9 \rightarrow 12$)	2013	1648.307	993.666	301.411	0.000	6515.770	423.525

1.4.2 Descriptive statistics

Table 1.2 reports information about parental monetary investment and HSP eligibility from age 5 to age 12 for the main sample. Median parental monetary investment rose from 754,398 VND reported in 2006 to 1,648,307 VND in real terms, outpacing inflation rates. This time trend in parental monetary investment is partially attributable to the maturity of the children in the sample. At the median value, the growth rate of educational investment is more rapid than above and below the median. By contrast, there exists a notable decline in parental monetary investment for entertaining and extra-activities as the child grows up, with the median value decreasing from about 94,429 VND in 2006 to almost zero in 2013 in real terms. This is consistent with a country-specific pattern that Vietnamese parents, similar to their counterparts in many East-Asian, Confucian-influenced countries, emphasize the importance and value of academic achievement by sacrificing other aspects of child development.

Table 1.3 summarizes information about family background factors for the main sample, disaggregated by the number of qualified criteria for HSP eligibility. Panel 1.3A lists control variables in my baseline specifications: wealth index, number of children, and parental education. Children in my sample, on average, have one sibling. The small number of siblings is partially due to the birth control policy that set limits on the maximum number of children to be two. Child fathers, on average, appear to have higher educational attainment than their spouses. Less than one-fifth of the sample lives in urban areas due to the oversampling of the poor in rural areas. Columns 2 and 3 of Table 1.3 show information about family background factors classified by number of satisfied criteria for HSP eligibility. Parents of children from families which are more likely qualified for HSP support tend to have lower educational attainment and from less wealthier families.

1.4.3 Capability gaps and parental monetary investment in Vietnam

This section documents evidence on the relationship between parental monetary investment and child capabilities in Vietnam during the 2000s and early 2010s.

Figure 1.1a and Figure 1.1b show the physical health gaps, measured by height-by-age z-scores, over family wealth quartiles. I focus on height because it is an important measure for child health and informative about economic and epidemiological conditions (Deaton, 2003, 2007, 2008; Pickles et al., 2007). As can be seen from Figure 1.1a, height-for-age of children from the low wealth families noticeably lags behind their wealthier peers, and the gaps persist throughout the whole childhood period. When I adjust for parental monetary investment[16], the gaps are slightly reduced but still considerable between the wealthiest children and the other groups.

Specifically, the gaps in height between youth from the first and the bottom quartiles of wealth can be reduced at most by 24.3 percentage points at age 5 and around 17-20 percentage points at the other ages. However, despite the reduced gaps, the ranking of child height closely tracks their ranks in the quartiles of wealth, even after taking into account the contribution of parental monetary investment. In other words, Figure 1.1b suggests that the children ranking on health closely track the ranks of their family wealth and

[16] I first regress the health measure at a specific age on parental monetary investment (log) and predict the regression residuals. I then rank children and construct percentiles by the residuals over time. I compute and plot the percentiles over time by wealth quartile. A similar procedure applies to cognitive skills and socio-emotional skills.

Table 1.3: Sample characteristics of children and their families

	Entire sample	Satisfied criteria for HSP		Difference
		>= 2 criteria	0-1 criteria	(2)-(3)
Panel A: Baseline variables				
Wealth index (continuous, 0-1)	0.615	0.684	0.529	0.155***
Number of children	2.045	1.873	2.166	-0.293***
Mother's education				
... Secondary school	0.538	0.472	0.586	-0.114***
... High school	0.163	0.093	0.212	-0.119***
Father's education				
... Secondary school	0.474	0.458	0.486	-0.028**
... High school	0.176	0.114	0.220	-0.106***
... College	0.097	0.049	0.131	-0.082***
Urban residence	0.191	0.033	0.301	-0.268***
Panel B: Additional variables				
Birth order	1.709	1.619	1.773	-0.154***
Child's age	6.663	5.735	7.310	-1.575***
Female	0.476	0.479	0.474	0.004***
Mother's age	32.762	31.342	33.578	-2.235***
... 2013	38.178	38.054	38.219	-0.165***
Household size	4.537	4.426	4.614	-0.187***
Number of child-year obs.	4780	1962	3584	
Number of children		1142		

remains rather constant throughout childhood. The presence of this obvious pattern is consistent with the view that long-term economic and sanitary conditions predominantly affect child physical health.

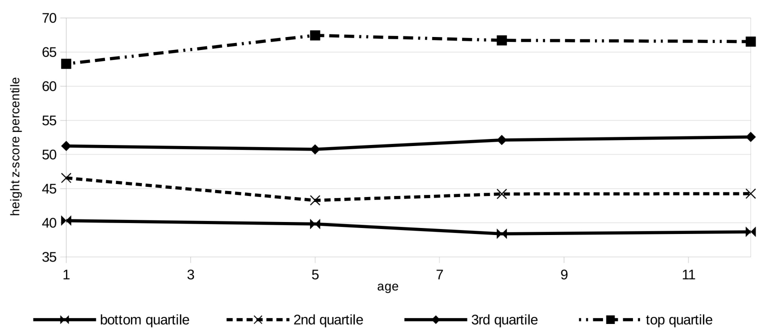
Somewhat different are the gaps in cognitive skills, measured by PPVT scores, for 5-12 year old Vietnamese children, which are shown in Figure 1.2a and Figure 1.2b. While the unadjusted PPVT gap in Figure 1.2b is quite similar to the persistent health gap,there is significant though not definitive evidence on the convergence in child cognitive skills which becomes stronger and stronger as children grow up when parental monetary investment is controlled for. In particular, the children from the second and third quartiles almost close the skill gap with children from the top quartile of wealth when they reach 12. This implies that the wealth-cognitive gradient observed in figure 1.2a seems to be due to differences in parental monetary investments across wealth quartiles.

However, there is also a cautionary tale concerning the performance of the most disadvantaged children from the bottom wealth quartile: their gap in cognitive skills compared to their counterparts remains significantly large even after controlling for the contribution of parental monetary investments. To sum up, three important observations emerge from my descriptive analysis. First, the gaps in physical health and cognitive skills (as measured by height and PPVT scores[17]) by family wealth open up before the children enter primary school and remain relatively large until age 12. This implies that the observed capability gaps are likely due to the wealth gaps. Second, parental monetary investment appears to play a significant role on the observed cognitive gaps. Third, when I adjust for the contribution of parental monetary investment, there is convergence in child health among SES groups. Moreover, the weak convergence in cognitive skills and health suggest that Vietnamese children from disadvantaged families have continued to lag behind their wealthier counterpart.

[17]Using the YLVN data, I follow a conventional practice and discuss the capability gaps in terms of units obtained by standardizing health measures and cognitive test scores to have within-sample zero mean and standard deviation equal to one. Cognitive measures are scores estimated from item response models (IRT) so that standard deviations serve as a cardinal metric for gaps in measures of skills, which, in turn, allows me to provide coherent information about the magnitude of the score gaps among different SES groups over time.

Figure 1.1: Development trajectory of child physical health

(a) Development trajectory of child physical health, unadjusted



(b) Development trajectory of child physical health, adjusted by monetary investment

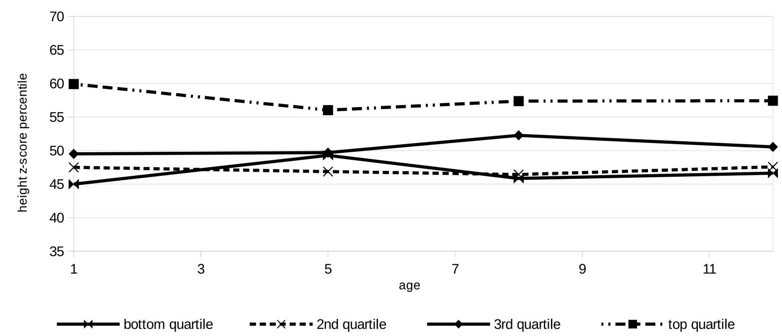
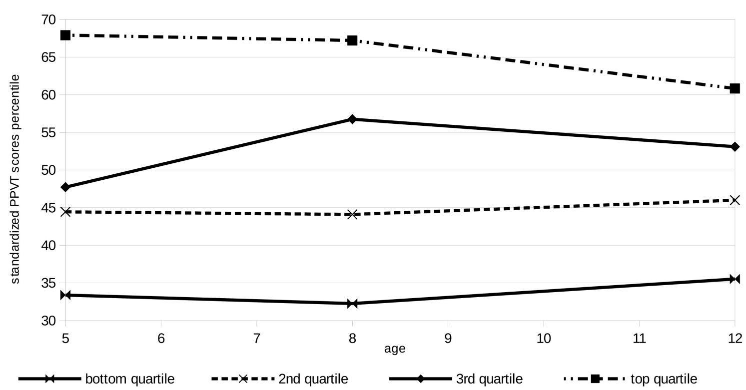
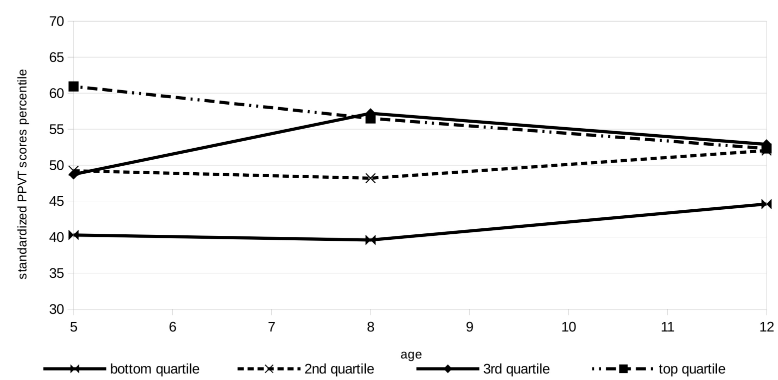


Figure 1.2: Development trajectory of child cognitive skills

(a) Development trajectory of child cognitive skills, unadjusted



(b) Development trajectory of child cognitive skills, adjusted by monetary investment



1.5 Estimation results

This section discusses the main results of my empirical analysis. Subsection 1.5.1 presents findings from the estimation of the measurement model. I focus on the information content of the measures and discuss the capability gaps by family wealth using the estimated latent capability factors. Subsection 1.5.2 is about the baseline Cobb-Douglas technology (1.4) with exogenous parental monetary investment. I then discuss the estimates of the Cobb-Douglas technology (1.4) with endogenous parental monetary investment in Subsection 1.5.3. Finally, Subsection 1.5.4 presents estimates for the translog technology (1.5) with endogenous parental monetary investment. I focus on two aspects: the complementarity between parental monetary investment and child capabilities, and the estimation of parental monetary investment decision.

1.5.1 The measurement system

1.5.1.1 The informational content of measures

This section discusses the informational content of measures and the importance of accounting for measurement errors when estimating the production technology. Following Cunha and Heckman (2008), the variance decomposition implied by the measurement system (1.2) is:

$$s_j^{ln\theta_{k,s}} = \frac{\alpha^2_{2,j,k,s}\text{var}(ln\theta_{k,s})}{\alpha^2_{2,j,k,s}\text{var}(ln\theta_{k,s}) + \text{var}(\mu_{j,k,s})} \tag{1.17}$$

I use a variety of measures to proxy for child capabilities, which differ from stage to stage. However, as emphasized in Section 3.1, I am able to normalize each capability k on the same normalizing measure in every stage s. The normalizing measures are height-for-age z-scores for health, PPVT test scores for cognitive skills, and self-efficacy for socio-emotional skills. This allows for comparisons of latent factors over times in a dynamic setting, as discussed in Agostinelli and Wiswall (2017).

Table 1.4 reports the shares of total variance associated with the latent factors (signal-to-noise ratios). These signal-to-noise ratios indicate the importance of correcting for measurement errors in estimating the production function of child capabilities. I find that the measures are highly informative about the latent variables and the informational content tends to be higher at the earlier stages of childhood. However, these measures are not inclusive and are imperfect manifestations of the underlying child capabilities. For example, height-for-age z-scores recorded at the second stage - the most informative measure among all measures, still has 2.73% of its total variance contaminated by measurement errors.

Cognitive skills and socio-emotional skills measures tend to be equally informative, with the signal-to-noise ratios of the richest and poorest measures within each stage differing by at most 52 percentage points (CDA and PPVT scores in the second stage). By contrast, the informational gaps among health measures are much greater, up to 73 percentage points (79.4% versus 6%) in the first stage. A large amount of the noise is associated with the caregiver's assessment of the child's general health. For example, the estimated latent health at the second stage accounts for approximately only 7% of the total variance of the caregivers' assessment about child health. Signal shares of latent health relative to total variance of that measure after stage 2 are much higher, suggesting that caregivers have better knowledge about child health as children grow up.

1.5.1.2 Parental monetary investment and child capabilities: using the latent factors

In Figure 1.3, I plot the mean of the health, cognitive skills, and socio-emotional skills factors against development stage for three long-term parental monetary investment percentiles: 25%, 50%, and 75%. The long-term parental monetary investment is the average of investments over twelve years. This is a counterpart to the descriptive exercise in Figure figure 1.1, except that now I use the latent factors and remove measurement error components.

Table 1.4: Signal and noise ratio of capability measures

	Stage 1 Age 0→1	Stage 2 Age 1→ 5	Stage 3 Age 5→ 8	Stage 4 Age 8→ 12
Measures of child's health				
Height-for-age z-score	79.426%	97.265%	73.075%	73.073%
Birth-weight z-score	6.193%			
BMI-for-age z-score	29.890%	69.765%	67.261%	58.448%
Child's health status	7.369%	27.631%	18.869%	15.963%
Child's height status		30.294%		
Measures of child's cognitive skills				
PPVT scores (Rasch)		79.096%	45.657%	24.083%
CDA scores (Rasch)		27.517%	46.284%	
Math scores (Rasch)				47.373%
EGRA scores (Rasch)			32.570%	
Reading scores (Rasch)				34.689%
Measures of child's socio-emotional skills				
Self-efficacy			48.895%	88.116%
Self-esteem			47.490%	22.773%
Positiveness			33.637%	

PPVT: Peabody Picture Vocabulary Test, EGRA: Reading comprehension test, CDA: Cognitive Development Assessment.

The figure reveals substantial and persistent differences in child capabilities across the investment distribution. The largest differences seem to be between those in the top 75% of the distribution and those below the 50% percentile. The health gap and socio-emotional gap increase substantially with age. By contrast, the gap in cognitive skills appear to decline over time, suggesting that monetary investments seems to be less effective in producing cognitive skills.

Figure 1.3 suggests that, for child's health and socio-emotional skills, parental monetary investments seem to be equally effective throughout childhood, while they are less effective in promoting cognitive skills as children grow up. There might be three explanations for the difference in investment effectiveness during childhood, which are not necessarily mutually exclusive. First, as I already analyze in Section 3, the productivity of investment depends on the malleability of child capability. Second, the parents who invest more in children might also have employed a more efficient production technology. For example, at the same level of investment, better educated parents are likely to invest more efficiently in their children health and cognitive skills. Allowing for a flexible production function with TFP dynamics and technological efficiency as I do is therefore important to distinguish between the two explanations. In the next section, I show that both of them are indeed the case at work in the context of Vietnam.

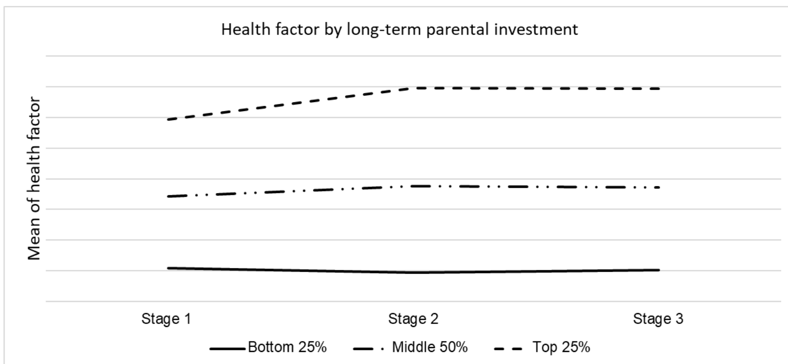
1.5.2 The Cobb-Douglas production technology with exogenous parental monetary investments

I start with the baseline production technology that characterizes the joint development of child health, cognitive skills and socio-emotional skills over four stages of of childhood. I here assume that monetary investment is exogenously determined and correct for measurement errors of latent capabilities by estimating the measurement system. The discussion focuses on contemporaneous effects of monetary investment. I postpone the discussion on the TFP dynamics and returns to scale to Section 1.5.3.4.

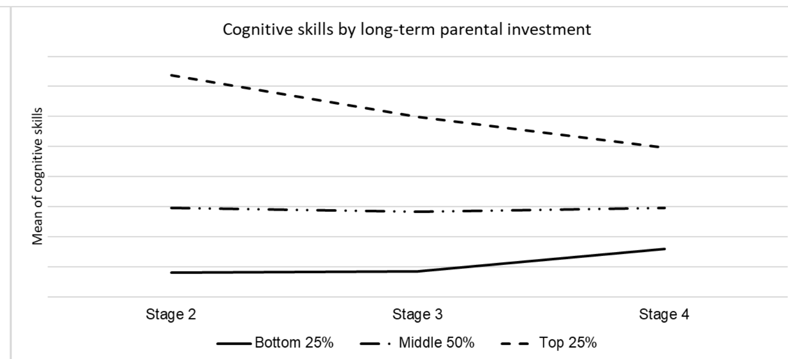
Table 1.5 reports estimates for the Cobb-Douglas production function in Equation (1.4). Parental monetary

Figure 1.3: Latent child capabilities by long-term parental monetary investment

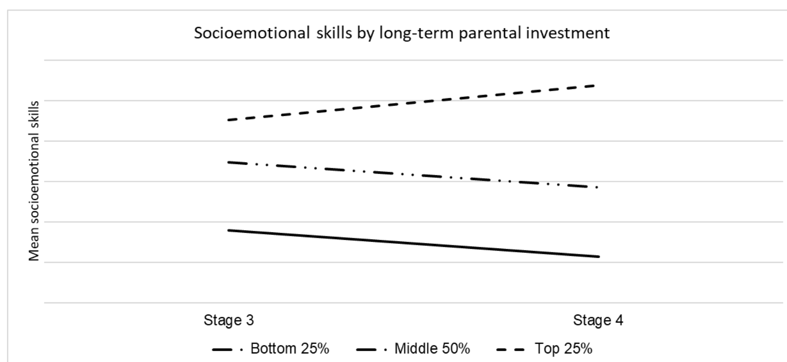
(a) Latent child health by long-term parental monetary investment



(b) Latent child cognitive skills by long-term parental monetary investment



(c) Latent child socio-emotional skills by long-term parental monetary investment



investment is most productive in earlier stages, specifically, before the child enters primary school, with respect to health and cognitive skills and during the first year in school for socio-emotional skills. Regardless of capabilities, the estimated magnitudes of investment productivity are quantitatively appreciable in the stages when they are statistically significant. Specifically, a 50% increase in parental monetary investment before the child goes to primary school raises child's latent health by about 4.8% ($\gamma_{H,I,2} \cong 0.077$) and cognitive skills by 10% ($\gamma_{C,I,2} \cong 0.202$) by the end of that stage. However, any increase in parental monetary investment made in the fourth stage does not induce statistically significant improvement in cognitive skills and socio-emotional skills. To illustrate how large a 50% increase in investment is, at the mean of parental monetary investments at age 5, a 50% increase would be equal to $32.5-$43 per year in 2006 prices (PPP), or 3.5% of the value of the average HSP one-time cash transfer. For comparison, individuals in the Young Lives sample have a monthly consumption per capita in 2006 of about $90-$120 (PPP). Yearly fees for public pre-schools in Vietnam in 2006 ranged from $6 to 120$ in urban areas and from $3 to $30 in rural areas (PPP), while public primary schools were free for all students.

There is evidence of an increasing self-productiveness of cognitive skills and health as the child grows up. In particular, health is highly self-productive after age five with the estimated parameters are close to unit in the third and fourth stages, $\gamma_{H,H,3} \cong 0.914$ and $\gamma_{H,H,4} \cong 0.875$. While the estimates for cognitive skills and health are as expected, there is no evidence for self-productiveness of socio-emotional skills. Specifically, socio-emotional skills in the third stage does not promote themselves in the fourth stage, with the coefficient $\gamma_{NC,4,3} \cong 0.026$ statistically insignificant, despite the fact that the measures of socio-emotional skills are highly informative across stages. This self-unproductiveness can be explained by the fact that the fourth stage of childhood in my analysis - from 8 to 12 years old, almost coincides with the pre-puberty period during which children are likely to experience significant mental changes.

Lastly, there is clear evidence of cross-productivity among all elements of child capability vector in all stages of development, not only indicating the crucial role of health but also the importance of modelling jointly the formation process of health, cognitive skills and socio-emotional skills. Panel A of Table 1.5 shows that child socio-emotional skills are productive in promoting health from age 9 to age 12, with $\gamma_{H,N,4} \cong 0.035$ (p-value = 0.037). Regarding cognitive skills, the impact of health is significant and appreciable in the earlier stages, before age 8, being $\gamma_{C,H,2} \cong 0.059$ and $\gamma_{C,H,3} \cong 0.061$. Similar to the long documented evidence in the literature, I find that socio-emotional skills promote the development of child cognitive skills ($\gamma_{C,N,4} \cong 0.032$) and vice versa (see the third row of Panel C of Table 1.5) (Cunha and Heckman, 2008; Cunha et al., 2010). Notably, Panel C shows that cognitive skills have a remarkable impact on child socio-emotional skills, of which the cross-productivity almost doubles from stage 3 to stage 4. A possible explanation for the positive cross-productivity between these capabilities is thatchildren with better cognitive skills may hold a stronger belief about their ability to pursue academic and later life endeavors, which are measures of socio-emotional skills (academic and job-search efficacy).

1.5.3 The Cobb-Douglas production technology with endogenous parental monetary investments

This section investigates the effects of parental monetary investment on child abilities, accounting for the endogeneity of investment. Throughout, I jointly estimate the production technology (1.4), the decision rules of parental monetary investments (1.7), the anchoring outcome equation (1.12), and the initial condition (1.13). I use simulated criteria for eligibility as an instrumental variable and allow the error components of Equations (1.4), (1.7) and (1.12) to be correlated. Correcting for the endogeneity of parental monetary investment has substantial effects on the magnitudes and statistical significance of estimated investment productivity in the production technology of the three capabilities. It turns out that ignoring investment endogeneity causes the estimated investment coefficients to be biased downward. I also discuss the results for the investment equation, which can be seen as the approximation of the parental monetary investment decision rule. I find that parents invest more on more highly capable children and are most responsive to cognitive skills compared to health and socio-emotional skills. It is also worth emphasizing that the instrumental variables, simulated HSP eligibility in three stages, are highly significant even after controlling for family wealth and a wide range of family background factors. I discuss this result in greater details in subsection (1.5.4.2).

Table 1.5: Exogenous parental monetary investment: contemporaneous productivity of investment, self- and cross-productivity of capabilities

	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12
Panel A. Health production function			
Investment	0.077***	0.019**	0.030**
	(0.014)	(0.008)	(0.013)
Previous stage capability			
... Health	0.807***	0.914***	0.875***
	(0.015)	(0.007)	(0.012)
... Cognitive skills		0.001	0.013
		(0.008)	(0.022)
... Socio-emotional skills			0.035**
			(0.017)
Panel B. Cognitive skills production function			
Investment	0.202***	0.063***	0.010
	(0.022)	(0.014)	(0.006)
Previous stage capability			
... Health	0.059**	0.061***	0.008
	(0.024)	(0.011)	(0.006)
... Cognitive skills		0.350***	0.578***
		(0.013)	(0.011)
... socio-emotional skills			0.032***
			(0.008)
Panel C. Socio-emotional skills production function			
Investment		0.056**	0.030
		(0.022)	(0.037)
Previous stage capability			
... Health		0.009	-0.006
		(0.019)	(0.033)
... Cognitive skills		0.083***	0.166**
		(0.021)	(0.062)
... Socio-emotional skills			0.026
			(0.048)

Table 1.6 presents the estimates of Equation (1.4) for health, cognitive skills and socio-emotional skills, respectively, accounting unobserved heterogeneity. I focus on the investment productivity and self- and cross-productivity of child capabilities.

1.5.3.1 Health production

Panel A of Table 1.6 shows the estimates for health over three stages. Returns to investment in child's health is U-shaped, with productivity being the lowest in the third stage but significantly higher in the second and fourth stages. The result is qualitatively similar to the previous case, in which parental monetary investment is assumed to be exogenously given. However, the estimated coefficients of investment productivity are substantially higher at all stages when I correct for investment endogeneity. In particular, a 50% increase in investment would induce about 5%, 3% and 3.8% increases in the latent health in the second, third and fourth stages, respectively. This result suggests that investment before the child goes to primary school (age $2 \rightarrow 5$) and during the puberty period (age $9 \rightarrow 12$) are particularly important in producing child physical health.

Consistent with the estimates reported in Table (1.5), health is increasingly self-productive over time, with the corresponding coefficients $\gamma_{H,H,2} \cong 0.804$, $\gamma_{H,H,3} \cong 0.907$, and $\gamma_{H,H,4} \cong 0.868$, suggesting that there could be fewer opportunities for external interventions to make an impact on health. However, this does not mean that recovery is not possible. Specifically, from age 9 to age 12 the slight increase in parental monetary investment productivity and decrease in self-productivity implies that health interventions at this stage may be more effective compared with interventions in stage 2.

I find that socio-emotional skills have a statistically significant effect on health from age 9 to age 12 ($\gamma_{H,N,4} \cong 0.030$), but cognitive skills do not. The result suggests that, in the context of Vietnam, interventions that improve child socio-emotional skills during puberty periods, including self-esteem and self-efficacy, may have impacts on health as well. This is also consistent with the evidence from the psychology and economics literature on the effects of sociobehavioural skills on healthy behaviour and physical health (Arigo et al., 2014; Arduini et al., 2017). For example, Arduini et al. (2017) find that, in the U.S, adolescent girls with low self-esteem incur a marked risk of developing eating disorders, which in turn pose severe harmful impacts on physical health.

1.5.3.2 Cognitive skills production

Panel B of Table 1.6 shows the estimates of the production function of cognitive skills. The results on self- and cross-productivity of child capabilities are qualitatively and quantitatively similar to those in Panel B of Table 1.5, while I observe substantial changes in estimated productivity of parental monetary investments.

First of all, parental monetary investment is highly productive in producing cognitive skills in all stages, but the productivity is decreasing substantially as the children grow up. The coefficients suggest that a 50% increase in parental monetary investment would improve child cognitive skills by 11%, 5.2% and 2.5% in the second, third and fourth stages, respectively. The impact of parental monetary investments on cognitive skills from age 9 to age 12 falls to almost one fourth of the impact from age 2 to age 5, indicating a significant slowdown in the productiveness of investment. Similar to health, I find that ignoring unobserved heterogeneity is likely to bias downward the estimates of investment productivity. Parental monetary investment still has a positive impact on child cognitive development up through age 12, although the impact is much smaller than in earlier stages.

Second, similar to the self-productivity of the health factor, I find that the self-productivity of cognitive skills is highly significant and increasing over time, being $\gamma_{C,C,3} \cong 0.339$ from age 5 to age 8 and almost doubled at age $8 \rightarrow 12$ with $\gamma_{C,C,4} \cong 0.565$. Notably, the evidence of cross-productivity from health and socio-emotional skills remains pronounced, as seen in Panel B of Table (1.5). The cross-productivity of health on cognitive skills in earlier stages is crucial, especially in the context of developing countries where malnutrition and illness - indicators of poor physical health - all have been demonstrated to have detrimental effects on child cognitive development (Lupien et al., 2000; McCoy et al., 2016; Walker et al., 2007). The findings also help to explain

Table 1.6: Endogenous investment: productivity of parental monetary investment, self- and cross-productivity of capabilities.

	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12
Panel A. Health production function			
Investment	0.098***	0.060***	0.070***
	(0.014)	(0.011)	(0.015)
Previous stage capability			
... Health	0.804***	0.907***	0.868***
	(0.015)	(0.007)	(0.012)
... Cognitive skills		-0.010	0.001
		(0.008)	(0.022)
... Socio-emotional skills			0.033**
			(0.017)
Panel B. Cognitive skills production function			
Investment	0.223***	0.104***	0.051***
	(0.023)	(0.016)	(0.010)
Previous stage capability			
... Health	0.056**	0.055***	0.001
	(0.024)	(0.011)	(0.006)
... Cognitive skills		0.339***	0.565***
		(0.013)	(0.011)
... socio-emotional skills			0.030***
			(0.009)
Panel C. socio-emotional skills production function			
Investment		0.097***	0.071*
		(0.024)	(0.037)
Previous stage capability			
... Health		0.003	-0.013
		(0.019)	(0.033)
... Cognitive skills		0.073***	0.154**
		(0.021)	(0.062)
... Socio-emotional skills		0.002	0.023
		(0.021)	(0.048)

why interventions such as the Jamaican program (Grantham-McGregor et al., 1991; Walker et al., 2011; Gertler et al., 2014) and the Colombian study (Attanasio et al., 2014; Attanasio et al., 2017), which target malnourished children through psychological stimulation and micronutrition supplements, deliver significant positive impacts on child cognitive skills.

1.5.3.3 Socio-emotional skills production function

Parental monetary investment is found to have significant positive effects on child socio-emotional skills from age 6 to age 12 as indicated in Panel C of Table 1.6. Similar to the production technology of health and cognitive skills, investment is relatively more productive in earlier stages, and correcting for its endogeneity has a substantial impact on the estimated investment productivity. Specifically, a 50% increase in parental monetary investments would induce a 5% increase in child's socio-emotional skills from age 6 to age 8 ($\gamma_{N,I,3} \cong 0.097$), and 3.6% increase from age 9 to age 12 ($\gamma_{N,I,4} \cong 0.071$). Moreover, investment is found to be statistically significant in the fourth stage when its endogeneity is taken into account. My results on the investment productivity over time differ from the findings of Cunha and Heckman (2008) and Cunha et al. (2010) in that they find that parental monetary investments are more productive in later stages of childhood. However, there are few reasons to expect the results should coincide given the different populations I investigate and the different measures of parental monetary investments.

With respect to the impact of current stocks of capabilities, I find that socio-emotional skills are not productive in producing themselves, pointing to the unstability of child's socio-emotional skills druing the puberty period, around age 9 to age 12[18]. However, cognitive skills have considerable positive effects on socio-emotional development and the cross-productivity level is doubled between stage 3 and stage 4, being $\gamma_{N,C,3} \cong 0.073$ and $\gamma_{N,C,4} \cong 0.154$, respectively. This finding is expected is and consistent with the evidence from the psychological and educational literature, which emphasizes that self-esteem and self-efficacy - two aspects of socio-emotional skills in this study - are highly correlated with academic ability (see, e.g., Bachman and O'malley, 1986; Purkey, 1970; Purkey and Novak, 1996; Calsyn and Kenny, 1977; Lawrence, 2006; Bandura et al., 1996). Moreover, in Panel B of Table 1.6, there is a strong evidence on the cross-productivity from socio-emotional skills to cognitive development with $\gamma_{C,N,4} \cong 0.03$ being highly significant. Combining these two results likely point to a long documented culture-specific fact that in East Asian cultural tradition, which is shared by Vietnamese families, scholastic ability is highly valued and leads to increases in self-esteem and self-improvement and vice versa (see, e.g., Spencer 1990; Spencer et al. 1991; Caplan et al. 1991; Kao, 1995, 2001; Kao and Thompson, 2003; Schneider and Lee, 1990).

1.5.3.4 Total productivity factors and returns to scale

This section discusses the importance of the TFP terms and of allowing for nonconstant returns to scale in the production technology. Table 1.7 presents the estimates of the TFP terms (in log) and other family background factors in two cases: the Cobb-Douglas production technology with (i) exogenous monetary investments, and (ii) with endogenous monetary investments. Three important observations emerge from the estimates in Table 1.7. First, there is pronounced evidence of decreasing returns to scale for the production functions of all capabilities in every stage. Indeed, with the exception of the returns to scale of the health technology in stage 4, the null hypothesis of constant returns to scale, that is $H_0 : \sum_\ell \gamma_{k,\ell,s} = 1$ with $k \in \{H, C, N\}$, $\ell = \{H, C, N, I\}$ for each $s = 2, 3, 4$, is rejected in all other cases. This is in marked contrast with the assumption of CRTS imposed in previous work (Cunha et al. (2010) and Attanasio et al. (2017a, 2017b)).

Second, the estimated TFP dynamics in the production technology are of considerable magnitude and statistically significant in most stages (exception of socio-emotional skills equation from age 9 to age 12). Moreover, the TFP estimates are highest from age 2 to age 5 for health and cognitive skills and substantially lower at later ages, with the health TFP from age 9 to age 12 falling to about one half of the level from age 2 to 5 and the corresponding figure for the cognitive TFP being one fourth. This feature - in combination with the

[18]This statistical insignificance may also simply imply imprecise estimates compared to the estimates of the health and cognitive skills production functions.

previous results about the declining productivity of parental monetary investments and the increasing self-productivity of child capabilities altogether - suggest that existing capabilities and investments are relatively more productive and efficient from age 2 to 5 than in later stages.

Third, among the family background factors that affect TFP, the most notable is caregiver's education, which has a positive impact on cognitive development from age 2 to age 8 and health development before children go to school - from age 2 to age 5. The negative effect of caregiver's education on health in the very first years in primary school is striking. One explanation comes from a country-specific fact that Vietnamese parents highly value academic achievement compared to the development of other aspects, such as health and socio-emotional skills. As they grow up, children from better off families - having highly educated parents, under greater pressure, are more likely to sacrifice health development for academic success. The estimated coefficient of sibling size is statistically significant and negative at stage 2 and stage 4, suggesting that being born in a larger family likely has negative impact on child health especially in early ages and during the puberty period. With limited family resources, as in most developing countries, a higher number of children likely has adverse effects on child development. In my context, I find that there are both direct effects of sibling size on child capabilities, and indirect effects though parental monetary investments.

1.5.4 The translog technology with endogenous parental monetary investments

1.5.4.1 The translog technology and complementarity between parental monetary investment and capabilities

Table 1.8 reports the parameter estimates for the translog technology of capability formation in Equation 1.5. I focus on the investment productivity and its implications for estimating the technology and policy interventions. I correct for the endogeneity of parental monetary investments as explained in subsection 1.3.4.4. I also show the matrix of correlations between investments terms and child capabilities in Table 1.9, emphasizing the importance of correcting for its endogeneity.

Overall, two important results emerge. First, comparisons with the estimates of the Cobb-Douglas production function with endogenous investment in Section (1.5.3) show that the patterns of self-productivity and cross-productivity are almost unchanged. As before, cognitive skills and health are found to be increasingly self-productive, and all capabilities are positively increasing in parental monetary investments. There are positive cross-productivity of health on cognitive skills and of cognitive skills on socio-emotional skills. Second, Panel B of Table 1.8 reports the coefficients of parental monetary investments and interaction terms from the parsimonious model. The most important finding is the complementarity between investments and capabilities in the health and cognitive skills production technology. For the technology of noncognitive skills, although parental monetary investments are productive *per se*, I find no evidence of complementarity between investments and capabilities.

Health. There is strong evidence of complementarity among investments, health (stage 2) and socio-emotional skills (stage 2 and stage 4), with coefficients being statistically at the 1% level of significance. This implies larger returns to investment for children who are healthier and have higher stocks of socio-emotional skills. This complementarity is crucial because it is a potential source of inequality. Starting at the same level of stock of health and socio-emotional skills, wealthier parents invest more in children and their stocks of health in the next stage would be higher. Moreover, at the same level of investments, the returns to investment for children with higher existing stock of health and socio-emotional skills are higher. The finding that the initial health of children from better off families (wealthier, highly educated parents), is in fact better and that parents invest more on more able children, as shown in Section 1.5.4.2, only reinforces this result.

Cognitive skills. Similar to the health technology function, I find significant evidence of complementarity between investments and capabilities, but on different dimensions and at different stages (stage 3 and stage 4). The returns to investment in cognitive skills are significantly higher for children with high existing cognitive skills, socio-emotional skills and better health. Combined with the results from Panel A of Table 1.8 that health has positive cross-effects on cognitive skills, this finding emphasizes the crucial role of health in child development and shows that accounting for the presence of health in the human capital development process is nontrivial and fundamental.

Table 1.7: TFP dynamics and returns to scale

	Exogenous Investment			Endogenous Investment		
	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12
Panel A. Health production function						
Caregiver education	0.042**	-0.004	0.022	0.035**	-0.017**	0.014
	(0.013)	(0.007)	(0.013)	(0.013)	(0.008)	(0.013)
Number of siblings	-0.044**	-0.002	-0.031**	-0.044**	-0.001	-0.029**
	(0.015)	(0.006)	(0.010)	(0.015)	(0.006)	(0.010)
Log TFP	0.181***	0.029	0.048	0.228***	0.073**	0.095*
	(0.046)	(0.021)	(0.039)	(0.047)	(0.023)	(0.040)
Returns to scale (*)	0.884***	0.934***	0.953***	0.976***	0.958***	0.938***
	(0.018	(0.01)	(0.027)	(0.027)	(0.013)	(0.037)
Test for CRTS (*) ($\chi^2(.)$)	41.60***	41.49***	3.09*	28.51***	13.94***	1.08
Panel B. Cognitive production function						
Caregiver education	0.154***	0.110***	0.006	0.148***	0.097***	-0.002
	(0.022)	(0.013)	(0.006)	(0.022)	(0.013)	(0.007)
Number of siblings	-0.012	-0.007	-0.002	-0.011	-0.005	-0.000
	(0.024)	(0.011)	(0.005)	(0.024)	(0.011)	(0.005)
Log TFP	0.159*	-0.175***	0.003	0.207**	-0.131***	0.050**
	(0.076)	(0.036)	(0.019)	(0.077)	(0.037)	(0.022)
Returns to scale	0.261***	0.474***	0.628***	0.329***	0.52***	0.65***
	(0.03)	(0.018)	(0.013)	(0.045)	(0.021)	(0.019)
Test for CRTS (*) ($\chi^2(1)$)	604.32***	892.93***	762.34***	567.87***	759.18***	624.03***
Panel C. Socio-emotional skills production function						
Caregiver education		0.015	-0.021		0.002	-0.029
		(0.021)	(0.036)		(0.021)	(0.036)
Number of siblings		-0.028	-0.001		-0.027	0.000
		(0.018)	(0.030)		(0.018)	(0.030)
Log TFP		0.062	0.073		0.106*	0.120
		(0.058)	(0.111)		(0.059)	(0.111)
Returns to scale		0.148***	0.216***		0.165***	0.136
		(0.029)	(0.077)		(0.034)	(0.107)
Test for CRTS (*) ($\chi^2(1)$)		866.40***	102.83***		797.07***	97.94***

(*) Test for constant returns to scale. For the Cobb-Douglas production technology: $RTS = \sum_{\ell} \gamma_{k,\ell,s} = 1$ with $k \in \{H, C, N\}$, $\ell = \{H, C, N, I\}$ for each $s = 2, 3, 4$.

Table 1.8: Translog production technology with TFP dynamics, nonconstant returns to scale

	Health			**Cognitive skills**			**Socio-emotional skills**	
	Stage 2	Stage 3	Stage 4	Stage 2	Stage 3	Stage 4	Stage 3	Stage 4
	Age 2→ 5	Age 6→ 8	Age 9→ 12	Age 2→ 5	Age 6→ 8	Age 9→ 12	Age 6→ 8	Age 9→ 12
Panel 1.8A: Productivity of existing stocks of capability and investment								
Existing stock of capability								
. . . Health	0.898***	0.889***	0.797***	0.052	0.089***	-0.009	0.007	0.005
	(0.027)	(0.008)	(0.018)	(0.046)	(0.013)	(0.009)	(0.021)	(0.038)
. . . Cognitive skills		-0.013	0.008		0.320***	0.596***	0.050**	0.102
		(0.009)	(0.032)		(0.015)	(0.014)	(0.025)	(0.067)
. . . Socio-emotional skills			0.115***			0.050**		0.019
			(0.031)			(0.017)		(0.062)
Investment	0.065***	0.040***	0.037**	0.178***	0.071***	0.037***	0.069***	0.024
	(0.015)	(0.010)	(0.017)	(0.026)	(0.016)	(0.009)	(0.024)	(0.034)
Panel 1.8B: Complementarity (substitution) between investment and capabilities								
... Investment × Health	0.045***	0.011	0.062***	0.028	0.028**	0.001	0.016	0.026
	(0.013)	(0.009)	(0.015)	(0.023)	(0.015)	(0.007)	(0.022)	(0.031)
... Investment × Cognition		0.002	0.035		0.025*	0.056***	-0.015	-0.036
		(0.009)	(0026)		(0.015)	(0.013)	(0.023)	(0.052)
... Investment ×Socioemotion			0.077***			0.050***		0.012
			(0.028)			(0.017)		(0.054)

Table 1.9: Correlation matrix: correlation of error terms between parental monetary investments and child capabilities.

	Health			**Cognitive skills**			**socio-emotional skills**	
	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12	Stage 2 Age 2→ 5	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12	Stage 3 Age 6→ 8	Stage 4 Age 9→ 12
Investment	-0.026***	-0.079***	-0.045***	-0.015***	-0.048***	-0.096***	-0.032***	-0.021***
	(0.006)	(0.020)	(0.011)	(0.004)	(0.012)	(0.023)	(0.008)	(0.005)
Interactions between investment and capabilities								
... Investment × Health	-0.019***		-0.037***		-0.041***			
	(0.005)		(0.009)		(0.010)			
... Investment × Cognition			-0.072***		-0.045***	-0.124***		
			(0.018)		(0.011)	(0.029)		
... Investment ×Socioemotion						-0.152***		
						(0.036)		
N	1131							

1.5.4.2 Estimates of the investment functions

As mentioned before, the investment equation (1.7), besides allowing me to account for endogeneity of investments, is also of independent interest. The estimated coefficients reported in Table 1.10 highlight the contributions of capabilities, HSP eligibility and other inputs to the parental monetary investment decision.

The first notable result is the positive impact of HSP eligibility on parental monetary investments, which is statistically significant. The fact that, being eligible for the HSP benefits increase the resources that parents provide to children is important because it shows that parents are likely to reinforce the potential HSP benefits. The HSP has been seen as a major, successful effort of the Vietnamese government in alleviating poverty, but its effects on child development have often been overlooked, both by policymakers and researchers. From a policy perspective this is a major conclusion that should encourage further investigation. Given that I use simulated eligibility instead of actual participation, the simulated eligibility would likely be only a lower bound of the effects of the actual participation in HSP.

Secondly, the coefficients of family wealth from age 2 to age 12 suggest that wealth plays a decisive role on the investment decision: a 0.22 point increase in the wealth index (recall that $0 \leq w_s \leq 1$) would induce about a 50% increase in parental monetary investment. Within any stage, the impact of wealth likely surpasses the combined effects of all other factors. This has a strong implication for policy in that, inequality in wealth, an indicator of socioeconomic status, perpeptuates inequality in human capital development in very early childhood, through its substantial effects on parental monetary investment. Without early intervention, the course is unlikely to be reversible.

Turning now to the roles of child own abilities, I find that parents invest more resources in children with higher levels of health, cognitive skills and socio-emotional skills. The impact of child cognition appears to be the strongest with $\gamma_{I,C,s} \cong 0.188$ for all s being double the impact of health $\gamma_{I,H,s} \cong 0.097$ and of socio-emotional skills $\gamma_{I,N,s} \cong 0.079$. Caregivers play an important role in investment decisions; since the effects of his/her educational level are significant in all stages and most pronounced in the early stage, before the child reaches 5 years old.

It is important to remember that the validity of the simulated HSP eligibility as instrumental variables for endogenous parental monetary investments relies on the following characteristics of the HSP. First, the HSP ties cash transfers to housing construction/improvement and allows beneficiaries to receive only a one-time unconditional cash transfer and/or a low interest loan. Eligibility for HSP benefits at stage s, a potential upward shift in the budget constraint and/or consumption smoothing, may induce (constrained) parents to increase child investment at stage $s' \geq s$. Second, the HSP changes its geographical coverage every five years over the period from 1999 to 2020, almost coinciding with the development stages I consider in this paper. This change would result in additional exogenous variation in parental monetary investments. Finally, small changes in the regulation on maximum/minimum financial and in-kind support also aid in identification.

One may also be concerned with the validity of my exclusion restrictions if HSP eligibility alters, and hence is correlated with, the time inputs which have been left in the error terms in my estimation. Previous studies have established the importance and productiveness of parental time input (Bernal and Keane, 2010, 2011; Del Boca et al. (2014)). Indeed, Bernal and Keane (2010, 2011) show that in the US context, eligibility for welfare benefits can alter parental time inputs, not just material inputs, because the welfare program (AFDC in their study) stipulates a certain time requirement for work activities to maintain benefits. That the HSP cash transfer and its bonus loan are classified as unconditional transfers, being eligible for HSP benefits in my context was neither conditioned on working status nor required the beneficiaries to work.

On the one hand, although the primary goal of HSP programme has been to improve the living conditions, there are still good reasons to believe that entitlement to HSP benefits, in particular the low-interest loan, can have impacts on family decision making such as labour supply and accumulation of assets (Asfaw et al., 2014). On the other hand, there is lack of rigorous evidence on the effects of unconditional cash transfers on labour supply in developing countries. The most recent review by Banerjee et al. (2017), re-analyzing seven RCTs of unconditional cash transfers in six countries, finds no systematic evidence that these programs discourage working incentives. Indeed, in my context, even if HSP eligibility changes individual working behaviour, it does not necessarily and obviously mean that parental time inputs would be altered accordingly.

Table 1.10: Approximation of parental monetary investment decision

	Cobb-Douglas production function
	Coefficients (Std. err)
HSP eligibility	0.064***
	(0.012)
Health	0.097***
	(0.015)
Cognitive skills	0.188***
	(0.023)
Socio-emotional skills	0.079**
	(0.037)
Caregiver's education	0.172***
	(0.016)
Wealth	2.238***
	(0.115)
Number of siblings	-0.062***
	(0.014)
Constant	-2.630***
	(0.090)
N	1131

1.6 Policy implications and conclusions

In this paper, I study the human capital development of Vietnamese children from age 1 to age 12 and focus on the role of parental monetary investments. I use data from the younger cohort of the Young Lives Study in Vietnam to estimate a dynamic, nonlinear factor model of production technology, following the models in Cunha et al. (2010) and Agostinelli and Wiswall (2016). I estimate the production functions of child health, cognitive skills, and socio-emotional skills jointly with the parental monetary investment equations. Health, cognitive skills, and socio-emotional skills are three crucial dimensions of individual capabilities and key determinants of later-life outcomes. These skills interact among themselves and with parental monetary investments in promoting children's human capital. This critical feature cannot be understood if each dimension of child capability is studied in isolation.

I allow for the possibility of complementarity (or substitutability) between monetary investments and child capability, and in the production technology for Hicks-neutral total factor productivity (TFP) dynamics and non-constant returns to scale. I correct for the possibility of endogenous parental monetary investments, taking advantage of the plausibly exogenous variation in parental monetary investments generated by the geographic- and time-variant eligibility for welfare benefits under the 1999-2020 National Target Programs in Vietnam. To avoid the issue of eligibility manipulation, I construct family-specific simulated eligibility for welfare programs and use these variables as instrumental variables instead of actual eligibility, which prove to be strongly correlated with parental monetary investments.

Given this model, I obtain a number of crucial results, which have not been available in the previous studies, especially in the context of developing countries. First, the three dimensions of child capabilities are cross-productive over the childhood. Specifically, low stocks of health in earlier ages cause permanent cognitive deficits at any later ages; low stocks of cognitive skills in earlier ages lead to lower levels of later socio-emotional skills; and during the pre-puberty period, low stocks of socio-emotional skills result in both health and cognitive deficits in later ages. While the results are consistent with what has been documented in previous studies, I am able to trace the effects and investigate their sources. The key policy implication from this exercise is that comprehensive policy interventions for disadvantaged children need to target and address

Table 1.11: Approximation of interaction between parental monetary investment and capabilities

	Translog production function			
Coefficients (Std. err)	Investment	Investment × Health	Investment × Cognitive skills	Investment × socio-emotional skills
HSP eligibility	0.027**	-0.025	-0.030**	-0.011
	(0.012)	(0.016)	(0.013)	(0.013)
HSP eligibility ×health	-0.018**	-0.199***	-0.042***	-0.002
	(0.010)	(0.013)	(0.011)	(0.013)
HSP eligibility × cognition	-0.075***	-0.070***	-0.149***	-0.007
	(0.016)	(0.019)	(0.016)	(0.024)
HSP eligibility ×socioemotion	-0.018	0.014	-0.016	-0.140***
	(0.033)	(0.041)	(0.026)	(0.021)
Health	0.100***	-0.432***	0.098***	0.012
	(0.015)	(0.030)	(0.021)	(0.021)
Cognitive skills	0.155***	0.207***	-0.350***	-0.057*
	(0.022)	(0.039)	(0.032)	(0.035)
Socio-emotional skills	0.036	-0.049	-0.043	-0.714***
	(0.039)	(0.057)	0.035	(0.029)
Caregiver's education	0.047***	0.024***	0.001	0.001
	(0.004)	(0.005)	(0.004)	(0.004)
Wealth	1.752***	0.021	0.116	-0.181
	(0.097)	(0.146)	(0.124)	(0.126)
Number of siblings	-0.087***	0.003	0.000	0.000
	(0.015)	(0.017)	(0.014)	(0.014)
Joint tests of significance of instruments by equation	$\chi^2(4) = 34.51$ $p-$value $= 0.000$	$\chi^2(4) = 284.7$ $p-$value $= 0.000$	$\chi^2(4) = 129.23$ $p-$value $= 0.000$	$\chi^2(4) = 47.01$ $p-$value $= 0.000$
Overall tests of significance of instruments	$\chi^2(16) = 498.36$ $p-$value $= 0.000$			
N	1131			

child health and cognitive deficits very early in life, i.e., before age 5, and also include measures to improve child socio-emotional skills starting from age 5 onward.

Second, parental monetary investments are crucial for promoting child health, cognitive skills, and socio-emotional skills. In particular, parental monetary investments are significantly productive at all ages for health (with productivity being U-shaped) and cognitive skills (with impact diminishing by age 8), and up to age 8 for socio-emotional skills. My finding that parental monetary investments continue promoting health development until age 12 is highly consistent with recent crucial findings from the Young Lives study on possibilities to correct for health deficits after the first 1000 days. However, my paper goes beyond those findings, showing clearly the mechanisms by which the health deficits can be recovered: parental monetary investments policy and interventions that include measures to address child's socio-emotional skills.

Third, I find that these three dimensions of child capabilities are significantly complementary to parental monetary investments at all ages in the production functions, implying that parental monetary investments are more productive on children already having higher stocks of capabilities. The complementarity between investments and health appear to be the strongest and most persistent over time, reinforcing the need to take into account child health in modelling the development of human capital. More importantly, as long as child human capital is a determinant of later-life outcomes, this complementarity perpeptuates and is a crucial source of socioeconomic inequalities. This calls for special attention to two aspects in the design of policy interventions: (i) on the one hand, the timing of intervention, (ii) on the other hand, the dimensions of child capabilities on which the intervention needs to be targeted. For example, it is clear from my analysis that interventions targeting children with health and cognitive deficits should also include measures to improve child socio-emotional skills, so that their full potentials on child capability development can be realized. In addition, the results of strong complementarities between child capabilities and parental monetary investment is particularly worrying in Vietnam, where economic growth has been accompanied by rising inequality of opportunities for children and youth and rising economic inequality in later life. The complementarity arises from age 1 and persists until age 12 (the latest period for which I have data), which underscores the need for sustaining policy interventions over all periods of childhood.

Finally, I also find that family wealth, which is informative about material wellbeing, plays an important and indispensable role both directly on the development of child capabilities and indirectly, through its substantial effects on parental monetary investments. In particular, at any age, children from wealthier families receive significantly more investments, even if they have similar stocks of initial capabilities. Material well-being has significantly positive direct effects *per se* on child health at all stages, up to age 8 for cognitive skills, and during age 9-12 for socio-emotional skills. This finding clarifies the two mechanisms (direct and indirect effects) through which policy interventions on family living conditions, such as the Housing Support Programmes in Vietnam, can bring about improvements on human capital. Policy interventions that target family material wellbeing can have substantial effects on child development indirectly through relaxing the budget constraints on parental monetary investments and directly on child capabilities, especially health and cognitive skills. This calls for a deeper analysis on the effects of welfare interventions such as the housing support programmes in Vietnam on child poverty and the intergenerational transmission of poverty.

This paper answers important and new questions on child human capital development in the context of a developing country. It also raises new problems, calling for a deeper analysis which will be addressed in my future work.For example, when the multidimensionality of child capabilities is taken into account, the tradeoff between early versus late parental monetary investments is unlikely a simple one. Not only does the productivity of parental monetary investments vary over time, but also the periods of highest productivity differ across capability dimensions.

Bibliography

[1] Agostinelli, F. and Wiswall, M. (2016a). 'Estimating the Technology of Children's Skill Formation.' NBER Working Paper 22442. 1

[2] Agostinelli, F. and Wiswall, M. (2016b). 'Identification of dynamic latent factor models: The implications of re-normalization in a model of child development.' NBER Working Paper 22441.

[3] Almond, D. (2006). 'Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population.' Journal of Political Economy, vol. 114(4):pages 672–712. 3

[4] Almond, D., Edlund, L., and Palme, M. (2009). 'Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden.' The Quarterly Journal of Economics, vol. 124(4):pages 1729–1772. 3

[5] Amemiya, T. (1985). Advanced econometrics. Harvard university press.

[6] Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. In Proceedings of the third Berkeley symposium on mathematical statistics and probability (Vol. 5, pp. 111-150).

[7] Arcidiacono, P. and Jones, J. B. (2003). 'Finite mixture distributions, sequential likelihood and the EM algorithm.' Econometrica, vol. 71(3):pages 933–946. 4

[8] Attanasio, O., Cunha, F., and Jervis, P. (2015). 'Parental Beliefs and Investments in Human Capital.' Mimeo UCL. 3.2

[9] Attanasio, O., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). 'Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial.' BMJ, vol. 349. 1, 8, 33

[10] Attanasio, O., Meghir, C., Nix, E., and Salvati, F. (2017). 'Human capital growth and poverty: Evidence from Ethiopia and Peru.' Review of Economic Dynamics, vol. 25:pages 234–259. 9

[11] Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2018). Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia.

[12] Attanasio, O., Meghir, C., & Nix, E. (2017). Human Capital Development and Parental Investment in India. Economic Growth Center, Yale University.

[13] Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). 'Remedying Education: Evidence from Two Randomized Experiments in India.' The Quarterly Journal of Economics, vol. 122(3):pages 1235–1264. 5

[14] Barcellos, S. H., Carvalho, L. S., and Lleras-Muney, A. (2014). 'Child Gender and parental monetary investments in India: Are Boys and Girls Treated Differently?' American Economic Journal: Applied Economics, vol. 6(1):pages 157–189. 5.3

[15] Behrman, J. R. (1996). 'The impact of health and nutrition on education.' The World Bank Research Observer, vol. 11(1):pages 23–37. 3

[16] Bernal, R. (2008). 'The effect of maternal employment and child care on children's cognitive development.' International Economic Review, vol. 49(4):pages 1173–1209. 1

[17] Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). 'Early life health interventions and academic achievement.' The American Economic Review, vol. 103(5):pages 1862–1891. 3, 5

[18] Black, M. M. (2003). 'Micronutrient deficiencies and cognitive functioning.' The Journal of Nutrition, vol. 133(11). 6

[19] Bleakley, H. (2007). 'Disease and development: evidence from hookworm eradication in the American South.' The Quarterly Journal of Economics, vol. 122(1). 3

[20] Bleakley, H. (2010). 'Malaria Eradication in the Americas: A retrospective Analysis of Childhood Exposure.' American Economic Journal: Applied Economics, vol. 2:pages 1–45. 6

[21] Bollen,KennethA. 1989.Structural EquationswithLatentVariables.New York:Wiley

[22] Bond, T. N. and Lang, K. (2013). 'The evolution of the black-white test score gap in grades K–3: The fragility of results.' Review of Economics and Statistics, vol. 95(5):pages 1468– 1479. 27

[23] Boneva, T. and Rauh, C. (2015). 'Parental Beliefs about Returns to Educational Investments: The Later the Better?' Tech. rep., Human Capital and Economic Opportunity Global Working Group. 3.2

[24] Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). 'Early Childhood Investments Substantially Boost Adult Health.' Science, vol. 343:pages 1478–1485. 1

[25] Carneiro, P., Lopez-Garcia, I., Salvanes, K., and Tominey, E. (2015). 'Intergenerational Mobility and the Timing of Parental Income.' Mimeo UCL. 3

[26] Case, A., Lubotsky, D., and Paxson, C. (2002). 'Economic Status and Health in Childhood: The Origins of the Gradient.' American Economic Review, vol. 92(5):pages 1308–1334. 1

[27] Chay, K. Y. and Greenstone, M. (2003). 'Air quality, infant mortality, and the Clean Air Act of 1970.' NBER Working Paper 10053. 3

[28] Chong, A., Cohen, I., Field, E., Nakasone, E., and Torero, M. (2016). 'Are there nutrientbased poverty traps? Evidence on iron deficiency and schooling attainment in Peru.' American Economic Journal: Applied Economics. 6

[29] Clark, S. E., Jukes, M. C., Njagi, J. K., Khasakhala, L., Cundill, B., Otido, J., Crudder, C., Estambale, B. B., and Brookera, S. (2008). 'Effect of intermittent preventive treatment of malaria on health and education in schoolchildren: a cluster-randomised, double-blind, placebo-controlled trial.' Lancet, vol. 372:pages 127–138. 6

[30] Cunha, F., Elo, I., and Culhane, J. (2013). 'Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation.' NBER Working Paper 19144. 1, 3.2

[31] Cunha, F. and Heckman, J. (2007). 'The Technology of Skill Formation.' American Economic Review, vol. 97(2):pages 31–47. 1

[32] Cunha, F. and Heckman, J. J. (2008). 'Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation.' Journal of Human Resources, vol. 43(4):pages 738–782. 1 4

[33] Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. Handbook of the Economics of Education, 1, 697-812.

[34] Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). 'Estimating the technology of cognitive and noncognitive skill formation.' Econometrica, vol. 78(3):pages 883 – 931. 1, 3, 3.1, 3.4, 3.4, 5.1, 5.4, 7, A.1

[35] Currie, J. (2009). 'Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development.' Journal of Economic Literature, vol. 47(1):pages 87–122. 27

[36] Currie, J. (2011). 'Inequality at birth: Some causes and consequences.' NBER Working Paper 16798. 1

[37] Currie, J. and Almond, D. (2011). 'Human capital development before age five.' Handbook of Labor Economics, vol. 4:pages 1315–1486. 1

[38] Currie, J. and Hyson, R. (1999). 'Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight.' American Economic Review, vol. 89(2):pages 245–250. 1

[39] Currie, J. and Neidell, M. (2005). 'Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?' The Quarterly Journal of Economics, vol. 120(3):pages 1003–1030. 3

[40] Currie, J., Neidell, M., and Schmieder, J. F. (2009). 'Air pollution and infant health: Lessons from New Jersey.' Journal of Health Economics, vol. 28(3):pages 688–703. 3

[41] Currie, J. and Stabile, M. (2003). 'Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children?' American Economic Review, vol. 93(5):pages 1813–1823. 1

[42] Del Boca, D., Flinn, C., and Wiswall, M. (2014). 'Household choices and child development.' The Review of Economic Studies, vol. 81(1):pages 137–185. 1, 3, 13, 34

[43] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). 'Maximum likelihood from incomplete data via the EM algorithm.' Journal of the Royal Statistical Society. Series B (Methodological), pages 1–38. 4

[44] Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. Child development, 82(1), 405-432.

[45] Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L., Martorell, R., Eming, M., Young, and Group", T. I. C. D. S. (2007). 'Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world.' The Lancet, vol. 369 (9557):pages 229–242. 4

[46] Fernald, L. C., Weber, A., Galasso, E., and Ratsifandrihamanana, L. (2011). 'Socioeconomic gradients and child development in a very low income population: evidence from Madagascar.' Developmental Science, vol. 14(4):pages 832–847. 1

[47] Field, E., Robles, O., and Torero, M. (2009). 'Iodine deficiency and schooling attainment in Tanzania.' American Economic Journal: Applied Economics, pages 140–169. 3

[48] Figlio, D., Guryan, J., Karbownik, K., and Roth, J. (2014). 'The effects of poor neonatal health on children's cognitive development.' The American Economic Review, vol. 104(12):pages 3921–3955. 1

[49] Florens, J. P., Heckman, J. J., Meghir, C., and Vytlacil, E. (2008). 'Identification of Treatment Effects Using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects.' Econometrica, vol. 76(5):pages 1191–1206. 17

[50] Galab, S., Kumar, S. V., Reddy, P. P., Singh, R., and Vennam, U. (2011). The Impact of Growth on Childhood Poverty in Andhra Pradesh: Initial Findings from India: Round 3 Survey Report. Young Lives, Department of International Development, University of Oxford. 10

[51] Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang-Lopez, S., and Grantham-McGregor, S. (2013). 'Labor Market Returns to Early Childhood Stimulation: A 20-year Followup to an Experimental Intervention in Jamaica.' NBER Working Paper 19185. 4

[52] Glewwe, P. and Jacoby, H. G. (1995). 'An economic analysis of delayed primary school enrollment in a low income country: the role of early childhood nutrition.' The Review of Economic Statistics, pages 156–169. 6

[53] Glewwe, P., Jacoby, H. G., and King, E. M. (2001). 'Early childhood nutrition and academic achievement: a longitudinal analysis.' Journal of Public Economics, vol. 81(3):pages 345– 368. 6

[54] Glewwe, P. and King, E. M. (2001). 'The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter?' The World Bank Economic Review, vol. 15(1):pages 81–113. 6

[55] Glewwe, P. and Miguel, E. (2008). 'The Impact of Child Health and Nutrition on Education in Less Developed Countries.' 5

[56] Glewwe, P., Krutikova, S., & Rolleston, C. (2017). Do schools reinforce or reduce learning gaps between advantaged and disadvantaged students? Evidence from Vietnam and Peru. Economic development and cultural change, 65(4), 699-739.

[57] Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. Multivariate Behavioral Research, 25(1), 33-39.

[58] Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., and Strupp, B. (2007). 'Developmental potential in the first 5 years for children in developing countries.' The Lancet, vol. 369(9555):pages 60 – 70. ISSN 0140-6736. 1

[59] Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., and Himes, J. H. (1991). 'Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: The Jamaican Study.' The Lancet, vol. 338(8758):pages 1–5. 5

[60] Griliches, Z. and Ringstad, V. (1970). 'Error-in-the-variables bias in nonlinear contexts.' Econometrica: Journal of the Econometric Society, pages 368–370. 3.4

[61] Gronau, R. (1974). 'Wage Comparisons–A Selectivity Bias.' Journal of Political Economy, vol. 82(6):pages 1119–1143. 17

[62] Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. Journal of Economic perspectives, 15(4), 57-67.

[63] Hausman, J. A., Newey, W. K., & Powell, J. L. (1995). Nonlinear errors in variables estimation of some Engel curves. Journal of Econometrics, 65(1), 205-233.

[64] Hamadani, J. D., Tofail, F., Huda, S. N., Alam, D. S., Ridout, D. A., Attanasio, O., and Grantham-McGregor, S. M. (2014). 'Cognitive deficit and poverty in the first 5 years of childhood in Bangladesh.' Pediatrics, vol. 134(4):pages e1001–e1008. 1

[65] Heckman, J. J. (1979). 'Sample Selection Bias as Specification Error.' Econometrica, vol. 47(1):pages 153–161. 17

[66] Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. Science, 312(5782), 1900-1902.

[67] Heckman, J. J. (2007). 'The economics, technology, and neuroscience of human capability formation.' Proceedings of the National Academy of Sciences, vol. 104(33):pages 13250– 13255. 1

[68] Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). 'The rate of return to the HighScope Perry Preschool Program.' Journal of Public Economics, vol. 94(1):pages 114–128. 5

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

[70] Heckman, J. J., Schennach, S., and Williams, B. (2010). 'Matching with error-laden covariates.' Unpublished manuscript, Department of Economics, University of Chicago. 1

[71] Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review, 103(6), 2052-86.

[72] Helmers, C., & Patnam, M. (2011). The formation and evolution of childhood skill acquisition: Evidence from India. Journal of Development Economics, 95(2), 252-266.

[73] Hoddinott, J., Maluccio, J., Behrman, J., Flores, R., and Martorell, R. (2008). 'Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults.' Lancet, vol. 371(9610). 5

[74] Hu, Y. and Schennach, S. M. (2008). 'Instrumental variable treatment of nonclassical measurement error models.' Econometrica, vol. 76(1):pages 195–216. 18

[75] Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success (No. w20749). National Bureau of Economic Research.

[76] Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. Journal of cognitive neuroscience, 16(8), 1412-1425.

[77] Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the National Academy of Sciences, 103(27), 10155-10162.

[78] Jayachandran, S. and Kuziemko, I. (2011). 'Why Do Mothers Breastfeed Girls Less Than Boys: Evidence and Implications for Child Health in India.' Quarterly Journal of Economics, vol. 126(3):pages 1485–1538. 3.1

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

[80] Kippler, M., Tofail, F., Hamadani, J. D., Gardner, R. M., Grantham-McGregor, S. M., Bottai, M., and Vahter, M. (2012). 'Early-life cadmium exposure and child development in 5- year-old girls and boys: A cohort study in rural Bangladesh.' Environmental Health Perspectives, vol. 120(10):pages 1462–1468. 6

[81] Kumra, N. (2008). 'An assessment of the Young Lives sampling approach in Andhra Pradesh, India.' 12

[82] Lu, C., Black, M. M., and Richter, L. M. (2016). 'Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level.' The Lancet Global Health, vol. 4(12):pages e916– e922. 2

[83] Lucas, A., Morley, R., and Cole, T. J. (1998). 'Randomised trial of early diet in preterm babies and later intelligence quotient.' BMJ, vol. 317(7171):pages 1481–1487. 5

[84] Ludwig, J. and Miller, D. L. (2007). 'Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design.' The Quarterly Journal of Economics, vol. 122(1):pages 159–208. 6

[85] Miguel, E. and Kremer, M. (2004). 'Worms: identifying impacts on education and health in the presence of treatment externalities.' Econometrica, vol. 72(1):pages 159–217. 5

[86] Murnane, R. J., Willett, J. B., Braatz, M. J., & Duhaldeborde, Y. (2001). Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. Economics of Education Review, 20(4), 311-320.

[87] Newey, W. K., Powell, J. L., and Vella, F. (1999). 'Nonparametric Estimation of Triangular Simultaneous Equations Models.' Econometrica, vol. 76(3):pages 565–603. 17

[88] Nielsen, E. R. (2015a). 'Achievement Gap Estimates and Deviations from Cardinal Comparability.' 27
Nielsen, E. R. (2015b). 'The Income-Achievement Gap and Adult Outcome Inequality.' 27

[89] Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N., and Grantham-McGregor, S. (2014). 'The socio-economic gradient of child development: Cross-sectional evidence from children 6-42 months in Bogota.' 1

[90] Sakti, H., Nokes, C., Hertanto, W., Hendratno, S., Hall, A., and Bundy, D. A. (1999). 'Evidence for an association between hookworm infection and cognitive function in Indonesian school children.' Tropical Medicine & International Health, vol. 4(5):pages 322– 334. 6

[91] Sazawal, S., Bentley, M., Black, R. E., Dhingra, P., George, S., and Bhan, M. K. (1996). 'Effect of zinc supplementation on observed activity in low socioeconomic Indian preschool children.' Pediatrics, vol. 98(6):pages 1132–1137. 5

[92] Schennach, S. M. (2004). 'Estimation of nonlinear models with measurement error.' Econometrica, vol. 72(1):pages 33–75. 18

[93] Todd, P. E. and Wolpin, K. I. (2007). 'The production of cognitive achievement in children: Home, school, and racial test score gaps.' Journal of Human Capital, vol. 1(1):pages 91– 136. 1

[94] Walker, S. P., Chang, S. M., Powell, C. A., and Grantham-McGregor, S. M. (2005). 'Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study.' The Lancet, vol. 366:pages 1804 – 1807. 1, 8, 33

Chapter 2

Marginal returns to upper secondary school in Indonesia: earnings and learning outcomes

Abstract

This paper estimates marginal returns to upper secondary school on the labour market and on learning outcomes in Indonesia. Using the longitudinal data from the Indonesian Family Life Survey 1997-2015, I document a substantial degree of heterogeneity in the returns to upper secondary school on the labour market. Marginal returns in earnings are found to be higher for individuals with characteristics that make them more likely to attend upper secondary school. In contrast, students with higher gains on learning outcomes are less likely to attend school. Moreover, students from disadvantaged backgrounds are not only less likely to go to upper secondary school but also have substantially lower marginal earnings returns. These findings suggest that universal upper secondary school expansion that successfully attracts low-resistant students who are currently not in upper secondary school may yield large pecuniary returns but are inequitable. Marginal expansions targeting disadvantaged students are likely to be both efficient and equitable than universal upper secondary school policies.

2.1 Introduction and motivation

Schooling expansion is at the heart of development policies in most low- and middle-income countries. When delivered properly, education improves earnings, employment, health and marriage outcomes. For societies, it strengthens institutions and socio-economic mobility as well as social cohesion through the generation of trust. In many countries, not only the speed but also the scope of expansion are historically unprecedented. Post-primary school is rapidly expanded in many developing countries, with some countries making upper secondary school universal or even compulsory. But much needs to be done. Achieving universal enrolment does not guarantee that schooling leads to higher learning outcomes and does not guarantee equality of labour market outcomes, especially for disadvantaged individuals (Crouch, 2006).

Despite enormous policy relevance, evidence about the marginal returns of upper secondary school expansion on the labour market in developing countries is scarce. Indeed, when evaluating the impact of secondary schooling expansion, the relevant quantities are the returns to students at the margins between enrolment or not, rather than the returns to the average student. A few exceptions are studies estimating both average and marginal retunrs to schooling in developing countries, such as Heckman and Li (2004) and Wang et al.

(2007) on returns to college in the Chinese labour market, and Carneiro et al. (2015) on returns to upper secondary school in the Indonesian labour market.

In this paper, I assess the marginal returns to upper secondary school on individual earnings in Indonesia - the fourth largest education system in the world (after only China, India and the United States). The goal is to better understand which individuals benefit most from schooling expansion to universal upper secondary school interventions, and the mechanisms through which schooling induces heterogenous effects on income. To do so, I estimate a semiparametric selection model of enrolment in upper secondary school using the marginal treatment effect model (MTE) (Heckman & Vytlacil, 2005, 2007). In this framework, returns to education are allowed to be heterogeneous across schooling choices and across individuals.

I report the returns to upper secondary school on the Indonesianlabour market and learning outcomes for a sample of 5209 Indonesian students aged 23 - 33 in 2015 using the Indonesian Family Life Survey 1997-2015. These cohorts are considered to be among the most relevant in emerging economies such as Indonesia.

My first finding is concerned with the existence of heterogeneous returns to upper secondary school on the Indonesian labour market, which is caused by both observed and unobserved characteristics. As for observed characteristics, students from wealthier families, having higher early cognitive skills and/or healthier are more likely to attend upper secondary school and receive higher wage returns, which points to the presence of selection on observed gains. The selection on individual unobserved characteristics reinforces this effect: students with unobserved characteristics that predispose them to upper secondary school benefit the most from schooling, whereas those who are least likely to attend benefit the least.As consequence, the returns to students currently in upper secondary school (the so-called average treatment effect on the treated, ATT) exceeds the returns of those opting out (the so-called average treatment effect on the untreated, ATU), with ATT being as high as 38 percent for each year of upper secondary school (statistically significant) and ATU being almost null (statistically insignificant). The upper secondary school expansion in Indonesia would attract students with lower earnings returns than the average returns of those currently in school. This pattern of selection on observed and unobserved pecuniary gains remains unchanged when modelling the schooling choice and earnings for different sub-samples of students. Because the OLS and conventional IV estimates commonly report only (local) average effects. The average estimates fail to reveal such important heterogeneity in returns to schooling.

Secondly, I show that the higher marginal returns for students, who are more likely to attend upper secondary school, are driven by lower earnings returns in the untreated state (without the qualification) and more homogeneous returns in the treated state (having the qualification). Moreover, these students are also more likely to come from advantaged backgrounds and have higher stocks of early cognitive and health capability. These results apply to the group of students who would change their schooling choices due to earnings gains unobserved to the analyst.

What, then, explains the pattern of selection into upper secondary school based on economic gains revealed in this paper? Why do students from advantaged backgrounds have higher marginal returns to upper secondary school, and are more responsive to marginal expansion of upper secondary schooling? To answer this question, I examine whether this economic inequality between the advantaged and disadvantaged students is a consequence of learning inequality. I investigate whether it is the case that students from disadvantaged backgrounds learn less than their better-off counterparts when they attend upper secondary school. This learning inequality would later be translated to earnings inequality as long as learning outcomes have positive impact on individual's earnings. Specifically, I investigate the returns heterogeneity to upper secondary school on student cognitive capability at adulthood (in a similar spirit to Cornelissen et al., 2019). The results reveal that students with a higher stock of early cognitive skills and coming from wealthier families also have higher adult cognitive ability, independently of schooling effects. However, attending upper secondary school not only promotes better cognition but can also (almost fully) compensate for early deficiency/disadvantages in those characteristics. In other words, students from disadvantaged backgrounds are likely to learn as much as those from advantageous backgrounds provided that they are in school and thus, learning inequality is unlikely the cause of the revealed inequality on the Indonesian labour market.

This paper contributes to the sparse research on heterogeneity in pecuniary returns to upper secondary school education in developing countries. I also contribute to the growing literature that estimates marginal treatment effects of education in different contexts, which has primarily focused on earnings returns to college in the developed countries (e.g., Carneiro et al. (2011) for the United States labour market, Balfe (2015) for the United Kingdom, Nybom (2017) for Sweden). Most of these studies depict substantial heterogeneity in marginal returns to schooling on the labour market, with marginal returns ranging from negative (some individuals incur loss from attending upper secondary school) to positive marginal returns which are much higher than the average returns. However, these studies do not enquire deeply into possible mechanisms causing heterogeneity, which is crucial in order to derive proper policy implications, as shown in this paper.

2.2 Estimating marginal returns to upper secondary school attendance

In this paper, I assess the marginal returns to upper secondary school on individual earnings and on learning outcomes in Indonesia - the fourth largest education system in the world. I estimate the marginal returns using the MTE framework (Björklund and Moffitt, 1987; Heckman, 1997; and Heckman and Vytlacil, 1999, 2005, 2007a; Carneiro et al., 2010, 2011, 2017).

Consider a simplified Becker-Mincer equation, $Y = \alpha + \rho S + \nu$, in which Y is the outcome of interest (log), S is schooling level, ρ is the rate of return to schooling, and α is the individual intercept. There are possibly two sources of estimation bias in the rate of returns ρ. The first source of estimation bias comes from the correlation between the unobserved disturbance and the schooling choice, i.e., $cov(\nu, S) \neq 0$. For example, if ν consists of individual ability which is positively correlated with schooling level S, i.e., $cov(\nu, S) > 0$, the OLS estimate of ρ will be upward biased. The second source of bias results from the correlation between schooling choice and returns to school, i.e., $cov(\rho, S) \neq 0$, which is termed essential heterogeneity. In this case, ρ is a random variable which is known by students and/or parents while unobserved by the analyst.

The MTE in this framework has several useful features. First, it provides the role of a function that is invariant to the choice of instrumental variables. Second, it has an attractive economic interpretation as the willingness to pay parameter for persons at the margins of indifference between selecting in school or not. Third, all conventional treatment parameters considered in the recent literature can be expressed as different weighted averages of the marginal treatment effects, such as the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the local average treatment effect (LATE). Using the method of local instrumental variables (LIV), the MTE can be identified and estimated under the standard IV assumptions of conditional independence and monotonicity (see Vytlacil 2002; Heckman 2010).

2.2.1 Defining the marginal returns to upper secondary school attendance

Potential and observed outcomes

In this section I follow Heckman and Vytlacil (2005, 2007a), Carneiro et al. (2010, 2011, 2017), and Brinch et al. (2017) and present the MTE approach that will be used to evaluate the existence and patterns of heterogeneous returns to upper secondary school attendance. The MTE framework can be seen as a generalized version of the Roy model (1951)

To start with, let Y_1 and Y_0 the potential outcomes for schooling levels “0” and level “1” respectively. The potential outcome Y_s, $s = \{0, 1\}$, is a function of control variables $\boldsymbol{X}$ (e.g., early family SES, community-level infrastructure, student age, religion) and early cognitive skills Θ_1 (basic literacy and numeracy, and abstract reasoning):

$$Y_s = \mu_s(\boldsymbol{X}, \Theta_1) + U_s, \;\; s = \{0, 1\} \tag{2.1}$$

where s indicates the schooling status and U_s is stochastic shock to the potential outcome Y_s.

The realized outcome Y is linked to the potential outcomes and schooling choices by:

$$\begin{aligned} Y &= (1 - S)Y_0 + SY_1 \\ &= Y_0 + S(Y_1 - Y_0) \end{aligned} \tag{2.2}$$

Equation (2.1) and Equation (2.2) imply that the effects of schooling on outcome Y can be written as the difference in potential outcomes in two states $S = 0/1$:

$$\triangle Y = Y_1 - Y_0 = \mu_1(\boldsymbol{X}, \Theta_1) - \mu_0(\boldsymbol{X}, \Theta_1) + U_1 - U_0. \tag{2.3}$$

At the individual level, the schooling effects on outcome Y varies with stocks of early cognitive abilities Θ_1, observed characteristics $\boldsymbol{X}$ and idiosyncratic shocks to potential outcomes (U_1, U_0). Equation (2.3) also implies that the net benefits of going to schooling level "1" can be partitioned into two parts: the observed part - $\triangle\mu(\boldsymbol{X}, \Theta_1) = \mu_1(\boldsymbol{X}, \Theta_1) - \mu_0(\boldsymbol{X}, \Theta_1)$, and the unobserved part $\triangle U_s = U_1 - U_0$.

Schooling choice

The schooling choice is motivated by the schooling returns on outcome Y, that is

$$S = \begin{cases} 1 & \text{if } Y_1 - Y_0 \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

that is individuals select into the schooling level "1" if the net expected returns to schooling is nonnegative.

Following Heckman and Vytlacil (2000, 2005, 2007a), I write a latent variable model that captures this decision rule as:

$$\begin{aligned} S^* &= \boldsymbol{Z}\gamma - V \\ S &= 1\{S^* \geq 0\} \end{aligned} \tag{2.4}$$

where the vector $\boldsymbol{Z} = (\boldsymbol{X}, \Theta_1, Z_+)$ includes the same controls as in Equation (2.2) $(\boldsymbol{X}, \Theta)$ and instrumental variables Z_+ excluded from the potential outcomes equation. Conditional on $(\boldsymbol{X}, \Theta_1)$, Z_+ affects schooling choices but not the realize outcomes, and thus, is uncorrelated with (U_1, U_0). Note that the unobserved shocks V enter the schooling choice equation with negative sign and reflect the unobserved factors that make individuals less likely to attend school. Following Cornelissen et al. (2016, 2018) I call V unobserved resistance or distaste for upper secondary school attendance. The higher is the value of V, the less likely is the student to attend upper secondary school.

Following the custom in the MTE literature, the schooling effects $\triangle Y$ can be traced out along the quantiles of the distribution $\mathcal{V}$ of the unobserved resistance V rather than its absolute values. Equation (2.4) can be transformed and rewritten as :

$$\boldsymbol{Z}\gamma - V \geq 0 \Leftrightarrow \boldsymbol{Z}\gamma \geq V \Leftrightarrow P(\boldsymbol{Z}) \equiv Pr(S = 1|\boldsymbol{Z}) = F_V(\boldsymbol{Z}\gamma) \geq F_V(V) \equiv \mathcal{V}$$

in which F is the cummulative distribution function of V, $P(\boldsymbol{Z})$ is the propensity score, i.e., the probability that a student with characteristics $(\boldsymbol{X}, Z_+)$ and early cognitive abilities Θ_1 will attend upper secondary school. $F_V(V)$ represents the quantiles of the distribution of distaste/resistance to upper secondary school V.

Model assumptions

Below, I summarize the assumptions about the random variables in Equation (2.1) and Equation (2.4), following the analysis of Heckman and Vytlacil (1999, 2001a, 2005), Carneiro et al. (2011), and Brinch et al. (2017).

Assumption 1. *The variables Z_+ induce variation in the propensity scores $P(\boldsymbol{Z})$ after controlling for $(\boldsymbol{X}, \Theta_1)$ in the schooling choice equation.*

For example, if distance to nearest upper secondary school is taken as an instrumental variable, the assumption requires that this distance influences schooling choices, after controlling for student early cognitive ability, family background factors, and community characteristics.

Assumption 2. *(V, U_0, U_1) is independent of Z_+, conditional on $(\boldsymbol{X}, \Theta_1)$.*

This assumption requires that the instrumental variables are as good as randomly assigned, conditional on $(\boldsymbol{X}, \Theta_1)$._

Assumption 3. $E(Y_s|V, \boldsymbol{X} = x, \Theta_1 = \theta_1) = \mu_s(\boldsymbol{X}, \Theta_1) + E(U_s|V), \quad s = 0, 1$.

The assumption means that the net unobserved gains $\triangle U = U_1 - U_0$ as a function of resistance to school V is independent of characteristics $\boldsymbol{X}$ and early cognitive ability Θ_1. This assumption is weaker than the additive separability between S and $(\boldsymbol{X}, \Theta_1)$ because it allows the treatment effects to vary by $(\boldsymbol{X}, \Theta_1)$ and V, although not by their interaction (Brinch et al., 2017).

Definition of marginal treatment effect (MTE)

The MTE measures the returns from attending upper secondary schooling for student with observed covariates $(\boldsymbol{X})$, early cognitive skills (Θ_1), and located at the v-th quantile of the $\mathcal{V}$ distribution (or those with propensity score of upper secondary school enrolment $P(\boldsymbol{Z})$ being equal to p), and is given as follows:

$$\begin{aligned} MTE(x, \theta_1, p) \equiv MTE(x, \theta_1, v) &= E(\triangle Y|\boldsymbol{X} = x, \Theta_1 = \theta_1, \mathcal{V} = v) \\ &= \triangle\mu(x, \theta_1) + E\left(\triangle U|\boldsymbol{X} = x, \Theta_1 = \theta_1, \mathcal{V} = v\right) \end{aligned} \tag{2.5}$$

Equation 2.5 means the MTE can be traced out within the support of propensity scores $P(\boldsymbol{Z})$ *conditional* on $(\boldsymbol{X}, \Theta_1)$. Brinch et al. (2017) show that Assumption 3 is sufficient for the separability of the MTE, i.e., the marginal returns to schooling (MTE) is additively separable into a unobserved and observed part:

$$MTE(x, \theta_1, v) = \underbrace{\triangle\mu(x, \theta_1)}_{\text{observed}} + \underbrace{E(\triangle U|\mathcal{V} = v)}_{\text{unobserved}} \tag{2.6}$$

In other words, Assumption 3, which implies the independence between $\triangle U$ and $(\boldsymbol{X}, \Theta_1)$, makes it possible to estimate the MTE over the *unconditional* support of $P(\boldsymbol{Z})$ instead of the conditional support of $P(\boldsymbol{Z})$. The marginal treatment effect is a function in which the constant is the treatment effect due to characteristics $\boldsymbol{X}$ and individual early cognitive ability Θ_1 and the slope, $E(\triangle U|\mathcal{V} = v)$, varies with individual's resistance to school but does not depend on $(\boldsymbol{X}, \Theta_1)$. This function is increasing (decreasing) in v if individuals who have high level of "distaste", i.e., high value of v, have higher (lower) returns to school.

Cognitive skills in early ages and at adulthood

In this paper, I exploit the availability of multiple cognitive tests scores prior to the upper secondary school entrance and at adulthood to extract information about student cognitive abilities. As already emphasized, I control for individual early cognitive ability in both the choice and outcome equations in estimating the marginal returns to school on the labour market and on learning outcomes in later life.

Let $T_{k,\tau}$ denote an individual's score on k-th test at period τ with $\tau = 1, 2$. I assume thacorresponding to age $7 - 14$ (prior to upper secondary school enrolment) and age 23-33 (adulthood), respectively. Assume that $T_{k,\tau}$ are finite. Thus, $T_{k,\tau}$ can be expressed as:

$$T_{k,\tau} = \gamma^T_{k,\tau} + ln\Theta_\tau \alpha^T_{k,\tau} + \epsilon^T_{k,\tau}, \ k = 1, ..., K; \tau = 1, 2 \tag{2.7}$$

in which the $\alpha^T_{k,\tau}$ are "factor loadings" that map the cognitive factor at period τ into test score $T_{k,\tau}$, $\epsilon^T_{k,\tau}$ are mutually independent and serially independent over time, $\epsilon^T_{k,\tau} \perp\!\!\!\perp (U_0, U_1, V)$ and $\epsilon^T_{k,\tau} \perp\!\!\!\perp (\boldsymbol{X}, \Theta_\tau)$. To set the location and scale of Θ_τ, I normalize $\alpha^T_{1,\tau} = 1$, so that $T_{1,\tau}$ is the anchoring measure, and $E(ln\Theta_\tau) = 0$ in each period τ. Moreover, to enable comparison between Θ_1 and Θ_2 in this dynamic settings, the anchoring measures $T_{1,\tau}$ are test scores of the same test (Agostinelli and Wiswall, 2016, 2018). Modelling test scores as in Equation (2.7) recognizes that they are manifestation of unobserved latent ability and contaminated by measurement errors[1].

While the MTE model does not require test scores for the identification of returns to schooling, the availability of test scores at $\tau = 1, 2$ offers several advantages. First, multiple test scores at adulthood allow me to investigate the effects of schooling on learning outcomes - an important mechanism through which education affects labour market outcomes. As argued by Glewwe (2002), more can be learnt from investigating the role of cognitive skills and its interaction with schooling on generating labour market outcomes rather than the schooling-wages relationship. In the context of developing countries, the interrelationship between the three variables - cognition, schooling, and earnings, is even more important because schooling does not automatically guarantee learning, which is often termed "the learning crisis". Second, multiple test scores in early childhood enable me extract information about the unobserved cognitive skills at early ages that affects both outcomes and schooling choices. Third, controlling for the early cognitive factor Θ_1 strengthens the validity of exclusion restrictions Z_+. The literature (see., e.g., Heckman et al., 2006b) has long acknowledged that most of the conventional instruments for schooling choices (e.g., nearest distance, sibling size, parental education and tuition fees) are correlated with individual cognitive skills, which also affect their later-life earnings. Regarding Assumption 2 of conditional independence, in the absence of Θ_1, the nearest distance and total number of accessible schools must be assumed to be independent of early cognitive ability left in the error terms (U_0, U_1). I discuss this point at length in subsection 2.3.4.

2.2.2 Estimating the marginal and average treatment effects

2.2.2.1 Estimating the marginal treatment effects

The main empirical analysis of this paper relies on a semiparametric estimation of the MTE, using the local instrumental variable (LIV) estimator as detailed in Heckman et al. (2006). In the following I summarize the main steps of the LIV estimator, following Heckman et al. (2006) and Carneiro et al. (2011, 2015). The idea is to rewrite the MTE, originally a function of $(\boldsymbol{X}, \Theta_1)$ and $\mathcal{V}$, as a function of $(\boldsymbol{X}, \Theta_1)$ and $P(\boldsymbol{Z})$, which are all observed and consistently estimated from data. For simplicity of notation, I assume that the choice and outcome equations are linearly separable in $\boldsymbol{X}$ and Θ_1, that is, $Y_s = \boldsymbol{X}\beta_s + \Theta_1\alpha_s + U_s$ with $s = 0, 1$, only in this section. As the result, the realized outcome Y equation in (2.2) is rewritten as:

$$Y = \boldsymbol{X}\beta_0 + \Theta_1\alpha_0 + S\left(\boldsymbol{X}(\beta_1 - \beta_0) + \Theta_1(\alpha_1 - \alpha_0) + U_1 - U_0\right). \tag{2.8}$$

In the empirical analysis presented below, I will instead allow for very flexible interactions between individuals' early capabilities and family backgrounds. The arguments with respect to the MTE estimation remain unchanged.

I exploit the fact that the model presented in Section 2.2.1 allows me to write the realized outcome in (2.8) as a function of the explanators $(\boldsymbol{X}, \Theta_1)$ and the propensity scores $P(\boldsymbol{Z}) = E\left(S = 1|\boldsymbol{Z}\right)$ (Heckman et al., 2006; Carneiro et al., 2011; Brinch et al., 2017):

$$\begin{aligned} E\left(\triangle Y|\boldsymbol{X} = x, \Theta_1 = \theta_1, P(\boldsymbol{Z}) = p\right) = \quad & \boldsymbol{X}\beta_0 + \Theta_1\alpha_0 + p\left[\boldsymbol{X}(\beta_1 - \beta_0) + \Theta_1(\alpha_1 - \alpha_0)\right] \\ & + \underbrace{pE\left(\triangle U|P(\boldsymbol{Z}) = p\right)}_{K(p)}, \end{aligned} \tag{2.9}$$

in which $K(p)$ is a function of propensity scores. Taking the first derivative of Equation (2.9) with respect to p produces the MTE evaluated at $\mathcal{V} = p$, $\boldsymbol{X} = x$, and $\Theta_1 = \theta_1$ (Heckman et al., 2006; Carneiro et al., 2011):

$$\begin{aligned} MTE(x, \theta_1, v) &= x(\beta_1 - \beta_0) + \theta_1(\alpha_1 - \alpha_0) + E(\triangle U|\mathcal{V} = p) \\ \frac{\partial E(Y|\boldsymbol{X}=x, \Theta_1=\theta_1, P(\boldsymbol{Z})=p)}{\partial p} &= x(\beta_1 - \beta_0) + \theta_1(\alpha_1 - \alpha_0) + \frac{\partial K(p)}{\partial p} \end{aligned}. \tag{2.10}$$

[1]I use the command sem in STATA version 14.2 (StataCorp, 2015) to estimate the latent factors (Θ_1, Θ_2)

Equation (2.10) suggests that estimating the MTE requires three components: (i) propensity scores $P(\boldsymbol{Z})$, (ii) the conditional expectation of Y, $E(Y|\boldsymbol{X}, \Theta_1, P(\boldsymbol{Z}))$, (iii) the first derivative of $E(Y|.)$ with respect to p, $k(p) \equiv k(v) = \partial K(p)/\partial p$. The estimation procedure consists of three steps, following closely the arguments above. The first step is estimating the schooling choice equation (2.4) and the propensity scores $P(\boldsymbol{Z})$, using a probit model $\hat{P}(\boldsymbol{Z}) = \Phi(\boldsymbol{Z}\hat{\gamma})$. The second step is to estimate the conditional expectation $E(Y|\boldsymbol{X}, \Theta_1, P(\boldsymbol{Z}))$ in Equation (2.9), in which the component $K(p)$ should be flexibly modelled. The more flexible $K(p)$, the more robust the estimated MTE. Finally, evaluating the derivative of $E(Y|\boldsymbol{X}, \Theta_1, P(\boldsymbol{Z}))$ with respect to p produces the MTE in Equation (2.10). I estimate the MTE model using the mtefe command written by Andresen (2018) in STATA. I describe the estimation procedure in details in Appendix A.1.

Equation (2.10) also suggests a simple test for the presence of heterogenous returns and selection on unobserved resistance to schooling that is to test whether $K(p)$ is a constant, or equivalently, the null hypothesis of $k(p) = 0$. Rejecting the null hypothesis implies the presence of heterogenous returns - the marginal returns to school vary with individual unobserved resistance to school. In the empirical estimation, I use this to test for the presence of unobserved heterogeneity and selection into school based on unobserved gains.

Finally, note that the true propensity score $P(\boldsymbol{Z})$ is not observed but is estimated in the first step using a probit model by $\hat{P}(\boldsymbol{Z})$, which clearly has estimation errors. This is true for the program evaluation studies relying on propensity scores. Therefore, one needs to adjust the estimated standard errors of the estimates to account for this estimation uncertainty (Abadie and Imbens, 2015). In the analysis below, I report confidence intervals which are estimated by bootstrapping. In each iteration, I reestimate every single step of the estimation procedure discussed above, from the probit model to the treatment effects estimation.

2.2.2.2 Estimating average treatment effects from the marginal treatment effects

Heckman and Vytlacil (1999, 2005, 2007) show that conventional average causal effect parameters, such as average treatment effect (ATE), average treatment effect on the untreated (ATU), and average treatment effect on the treated (ATT), can be constructed as weighted averages of the MTE curve. Specifically, these population average parameters are computed as follows:

$$\begin{array}{rcl} ATE(x, \theta_1) & = & \int MTE(x, \theta_1, v) f_{\mathcal{V}}(v) dv \\ ATU(x, \theta_1) & = & \int MTE(x, \theta_1, v) f_{\mathcal{V}}(v|S = 0) dv \\ ATT(x, \theta_1) & = & \int MTE(x, \theta_1, v) f_{\mathcal{V}}(v|S = 1) dv \end{array} \tag{2.11}$$

in which $ATE(x, \theta_1)$, $ATU(x, \theta_1)$, and $ATT(x, \theta_1)$ are conditional on $(X = x, \Theta_1 = \theta_1)$, and $f_{\mathcal{V}}(v)$ is the density of the quantiles $\mathcal{V}$ of the unobserved distaste for upper secondary school. The densities $f_{\mathcal{V}}(v)$, $f_{\mathcal{V}}(v|S = 0)$, and $f_{\mathcal{V}}(v|S = 1)$ are estimable weights[2] applied to corresponding (sub)populations of interest. I summarize the weights in the third column of Table 2.7 in Appendix.

In principle, these population average parameters can be evaluated at any value of $(\boldsymbol{X}, \Theta_1)$. However, following Cornelissen et al. (2016, 2018) I focus on the unconditional average parameters, that is, the ATE, ATU, and ATT are not only aggregated over the distribution of the unobserved resistance but also over the appropriate distributions of $(\boldsymbol{X}, \Theta_1)$. Provided that the MTE is additively separable, the weighted average of $(\boldsymbol{X}, \Theta_1)$ can be estimated separately using the weights in the fourth column of Table 2.7 in Appendix (see Cornelissen et al. (2016) for derivation of the covariate weights).

Another important average treatment-effect parameter is the local average treatment effect (LATE), which measures the average effects of schooling for individuals who would be induced to change schooling choice when the instrumental variables changes from $Z_+ = z_+$ to $Z_+ = \tilde{z}_+$. For any pairs $(z_+, \tilde{z}_+)$ such that $P(z_+) < P(\tilde{z}_+)$, these are individuals who would change from $S = 0$ to $S = 1$ and whose quantiles of the

[2] The original formulation of ATE, ATU, and ATT is derived by Heckman and Vytlacil (2005). I follow the representation of Carneiro et al. (2017) which is equivalent to the one in Heckman and Vytlacil (2005). In principle, these average parameters can be calculated at any value (x, θ_1) and the analyst needs integrating over all (X, Θ_1). That is, the weights should have been conditional on (x, θ_1) and written as $f_{\mathcal{V}}(v|X = x, \Theta_1 = \theta_1)$, $f_{\mathcal{V}}(v|X = x, \Theta_1 = \theta_1, S = 0)$, and $f_{\mathcal{V}}(v|X = x, \Theta_1 = \theta_1, S = 1)$. However, the additive separability of the MTE allows me to simplifies these conditional densities to be unconditional one - $f_{\mathcal{V}}(v)$, $f_{\mathcal{V}}(v|S = 0)$, and $f_{\mathcal{V}}(v|S = 1)$, respectively.

unobserved resistance $\mathcal{V}$ fall into the interval $(P(z_+), P(\tilde{z}_+))$. The LATE for a pair $(z_+, \tilde{z}_+)$ can be estimated as (Heckman and Vytlacil, 2005):

$$LATE(z_+, \tilde{z}_+) = \int MTE(x, \theta_1, v) f_{\mathcal{V}}(v | v_0 < \mathcal{V} < v_1) dv \tag{2.12}$$

with $v_0 = P(z_+)$ and $v_1 = P(\tilde{z}_+)$. It is important to emphasize that LATE in Equation (2.12) is defined by the instrumental variables used in the analysis and does not necessarily correspond to any (sub)population average parameters (Heckman, 1997; Deaton, 2009; Heckman and Urzua, 2010).

Finally, with continuous instruments, the traditional IV-2SLS parameter is a weighted average of all LATEs corresponding to all possible pairs $(z_+, \tilde{z}_+)$ (Angrist and Imbens, 1995), and therefore, can also be estimated by weighting the MTE curve[3]. In this paper, I use the IV-2SLS weights derived by Cornelissen et al. (2016) and summarize in Table 2.7 in Appendix. The estimation procedure of these weights is provided in Cornelissen et al. (2016) and Andresen (2018).

2.3 The data

2.3.1 Indonesian Family Life Survey

To analyze the marginal returns to upper secondary school on the Indonesian labour market and learning outcomes, I use data from four waves of the Indonesian Family Life Survey (IFLS) conducted in 1997, 2000, 2007 antd 2016. The IFLS is a household and community longitudinal study, conducted in 13 provinces and representing 83 percent of the Indonesian population. I analyze the cohort of individuals born between 1983 and 1992, aged 4-14 in the IFLS2 (1997/1998) and 23-33 in the IFLS5 (2015/2016). This cohort is particularly relevant for the Indonesian labour market, where the labour force is young and the economy is highly dynamic and growing rapidly. The IFLS study contains information on the highest level of completed schooling, individual capabilities prior to and after high school entrance, and annual earnings. The data also allow one to link individuals to their family background factors and community-level background during childhood.

2.3.2 Outcome variables: earnings and cognitive ability at adulthood

The IFLS study provides information on individual annual earnings, which include labour income from wage jobs and self-employment. I use income data of all those who reportedly worked between 2007 and 2014[4], because a sample of market-earnings earners would be more prone to sample selection bias in the context of developing countries (Glewwe, 2002). I deflate annual earnings to the base year in 2006.

Regarding the individual cognitive ability in adulthood, from the second wave in 1997 (IFLS2) the IFLS study administered a battery of cognitive tests (abstract reasoning, mathematics and language) to all individuals aged at least 7 to 24. The IFLS5 in 2015 retested all adults on mathematics skills and cognitive capacity (memory) when individuals in the main sample were from 23 to 33 years old. Therefore, measures of cognitive tests are available from both 1997 and 2015 for the target cohort born between 1983 and 1992. From these data, I extract information on individual cognitive abilities prior to upper secondary school entrance, which I use as an explanatory variable (Θ_1), and in adulthood, which is an outcome variable examined together with individual earnings.

The cognitive tests in the IFLS5 can be divided into two parts: (i) a set of cognitive tasks adapted for the Indonesian population[5] from the similar tests administered in the Health and Retirement Survey (HRS) in

[3] Heckman and Vytlacil (2005) derives the weights that apply to a general MTE model.

[4] As in other developing countries, the self-employed outnumber the earnings earners, accounting for about 61.56 percent (2007) to 51.11 percent (2016) of the total employment in Indonesia.

[5] These tests were extensively pretested in Indonesia and Mexico before the IFLS5 taking place. See Strauss et al. (2016, 2018) and Prindle and McArdle (2013).

the U.S; and (ii) an abridged version of the Ravens test. The HRS-adapted tests include: (i) a number series adaptive test, (ii) immediate and delayed word recall; and (iii) a task of serial subtraction of 7s from 100. The HRS-adapted tests and the Ravens test measure abstract reasoning ability and episodic memory (mental status intactness) (Ofstedal et al., 2005; McArdle et al., 2007; Strauss et al., 2016).

Both quantitative abstract reasoning and episodic memory are dimensions of fluid intelligence[6] which is the main dimension of cognition at adulthood, which I refer to as learning outcomes in this paper. Preliminary investigation using exploratory factor analysis reveals that this is indeed the case - the test scores of the three tests identify a single underlying factor. In the main analysis, I consider the test scores as manifestation measures of the unobserved fluid intelligence. This unobserved cognitive factor is identified and recovered using a measurement system widely used in the psychology literature and the economics literature on human capital development (for example, Bollen, 1989; Cunha and Heckman, 2008; Cunha et al., 2010; Agostinelli and Wiswall, 2018).

2.3.3 Explanatory variables: early cognitive ability and early health

Starting from the IFLS2 in 1997, all IFLS children older than age 7 were required to take cognitive assessments of their scholastic abilities (mathematics and language skills) as well as abstract reasoning. I use the scores from four tests, which were administered in 1997 and 2000. Specifically, in 1997 and 1998 (IFLS2), individuals between the ages of 7 and 24 received the mathematics and Indonesian language tests. The test items are drawn from the Indonesian National Achievement tests (EBTANAS). In 2000 (IFLS3), the tests were redesigned to cover skills in language, abstract reasoning and mathematics. I use only scores of tests taken by students aged 7-14 in the IFLS2 and IFLS3. This is to ensure that the students taking the tests were not yet enrolled in upper secondary school and, therefore, that their cognitive ability had not been affected by upper secondary school education.

The cognitive tests include multiple choice and open-ended questions. The IFLS study provides the information about children's answers to individual items of the test and whether these were correct[7]. Following the psychometrics and education literature, I use item-specific responses to construct children's test scores using a series of item response models (IRT)[8]. Prelimimary factory analysis reveals that these test scores are measures of a single latent factor. Using these test scores, I identify the distribution of the latent cognitive skills[9]. Similar to the latent cognitive skills at adulthood, this latent cognitive factor is separated out from measured cognitive abilities (test scores), from the effects of schooling levels at the test dates (as well as other background variables), and measurement errors.

The measure of each individual's health is based on a standardized evaluation aiming at determining each individual's physical health compared to peers of the same age. The evaluation is performed by trained health workers, who collect extensive measures of health status, including height, weight, head circumference, blood pressure, pulse, waist and hip circumference, hemoglobin level, and lung capacity. Based on those measures, the nurses then evaluate each individual physical health status on a 1 to 9 scale. In the analysis, I use the standardized scores of this evaluation within the IFLS population as the measure of early health.

[6] Cognition psychologists broadly classify cognition into fluid intelligence and crystalized intelligence (Horn and Cattell, 1966, 1967; McArdle et al., 2002). Abstract reasoning ability and episodic memory are dimensions of fluid intelligence, which is likely to peak at adolescence or in young adulthood. Crystalized intellect is accumulated through learning and tends to peak around 50 (Horn and Cattell, 1967; McArdle et al., 2002).

[7] The IFLS2 has information on the answer matrix of all children but does not provide answer keys to all test items for the mathematics and language tests. In a preliminary analysis, I produce answer keys to these test items. This data would be available upon request.

[8] For the foundational work on the theory of IRT models, see Rasch (1060), Birnbaum (1968), Wright and Stone (1979), Lord (1980). For recent advancement, see, e.g., Fischer and Molenaar (1995) and De Boeck and Wilson (2004). For the discussion on the advantages of using IRT models compared with raw scores (total sum of correct items) or the classical test model, see Samejima (1977).

[9] See other examples of the method in Carneiro and Heckman (2011), Heckman et al. (2015), Heckman et al. (2018a, 2018b).

2.3.4 Instrumental variables: distance to nearest upper secondary school and total number of accessible secondary schools

The two most commonly cited reasons for not attending school in Indonesia are unavailability of schools in the neighborhood and the financial burden of schooling attendance. In this paper I use two supply-side variables as instrumental variables for schooling choices: (i) the GPS distance from commune center to nearest upper secondary school accessible by community residents, (ii) the total number of upper secondary schools accessible by community residents. The exclusion restrictions are important for the identification of returns to upper secondary school. Specifically, they are both continuous variables rather than a simple dummy of whether a upper secondary school is available in the commune. While continuity of the IVs is not strictly required, this is the key feature that allow me to identify and estimate a wide range of different parameters on the returns to upper secondary school without parametric assumptions (Brinch et al., 2017).

Below I discuss the validity of these proposed instrumental variables for schooling choice. Distance to college has been used as an instrumental variable for college attendance in the literature by a number of studies. However, Heckman et al. (2006) argue that unless one controls for cognitive ability, the distance measure in the NLSY79 is an invalid instrument[10]. Indeed, several studies in the U.S context, using the NLSY79 data, have shown that distance to college at the college going age is correlated with a measure of cognitive ability (AFQT score) (Carneiro and Heckman, 2002; Cameron and Taber, 2004). In developing countries, long distance to upper secondary school may indicate disadvantaged local conditions and lower quality of schooling, which, in turn, affect individual's learning outcomes and their earnings. In this paper, I address this concern in two ways. I use available test scores to extract information on individual's unobserved cognitive skills and include this variable in both choice and outcomes equations, therefore, eliminate potential correlation between the nearest distance variable and unobserved parts of individual's earnings through individual's early cognitive ability. Moreover, I extract information about cognitive skills at adulthood and directly test whether the nearest distance has any effects on adult cognitive skills, conditioning on early cognition and other background factors.

Furthermore, it can be argued that the nearest distance to an upper secondary school might be correlated with both local and family socio-economic conditions. These individuals' background factors may also have effects on both schooling choices and earnings. In this paper, I control for a wide range of family background factors and community-level infrastructure availableduring childhood. The community infrastructure index is constructed similarly to the family wealth index and provides comprehensive information about infrastructure availability within the community- electricity, road, sewage system, piped water, and telecommunication. By including these variables as explanators in the choice and outcome equations, I avoid the possible correlation between the nearest distance and unobserved inputs.

Third, the nearest distance to school might be endogenous because individuals may strategically migrate to be closer to schools. The IFLS has a module in which parents were asked about the reasons for internal migration. One of the option was moving for education of other family members, i.e., including their children. Only three percent of IFLS respondents cited this as motivation for migration in the 2000s.

2.3.5 Analyzed sample

This paper uses the sample of children who were born between 1983 and 1992 in the IFLS data in 1997 and 2015. After removing those with missing information on individual backgrounds, early cognitive skills, earnings and cognitive skills at adulthood, the sample contains 5209 individuals. The sample is larger than that of other studies modelling the dynamics of schooling choices in Indonesia[11].Moreover, the age range of individuals in this paper (aged 22-32 in 2015) is much narrower, more relevant for a dynamic, emerging economy such as the Indonesian economy. Both these two features constitute advantages with respect to other studies in Indonesia.

[10] See Card (1995), Kane and Rouse (1995), Kling (2001), Currie and Moretti (2003), Cameron and Taber (2004), Carneiro and Heckman (2011), Carneiro et al. (2015).

[11] See Carneiro et al. (2015), Duflo (1990).

Schooling variable	**Equations**		
S=1 for highschool participants, S=0 otherwise	**Measurement**	**Choice**	**Outcomes**
Outcome variables			
- Annual earnings			x
- Adult cognitive ability (latent)			x
Observed covariates			
Early cognitive ability (latent)		x	x
Early health status		x	x
Family wealth index before age 12		x	x
Age		x	x
Gender (male =1)		x	x
Family size		x	x
Community-level wealth index before age 12		x	x
Measures of cognitive ability			
- Math test	x		
- Indonesian test	x		
- Logic test	x		
- Adaptive number series	x		
- Immediate and delayed word recall	x		
- Ravens test	x		
Instrumental variables			
GPS distance to the nearest upper secondary school**		x	
Total number of accessible upper secondary schools**		x	

Table 2.1: Dependent and explanatory variables, instrumental variables and measurement variables

Notes:

* : long-term SES index obtained by averaging dummies for family durable assets, including family size.

**: measured about 3-7 years before the enrolment decision into upper secondary school were made.

I estimate the model with the base group of individuals who had not graduated from upper secondary school (having no school up to having some years of upper secondary school) versus those graduating from upper secondary school or higher. This is the definition of upper secondary school attendance used in Carneiro et al. (2017). I present results for the sample that pools males and females, but includes a dummy variable to control for mean differences. Given that this definition of upper secondary school is arbitrary and the estimated parameters are sensitive to the base group, I proceed to estimate the model using two different base groups: (i) the individuals who graduated from primary school but did not attend upper secondary school graduates; (ii) the individuals who attended lower secondary school but did not attend upper secondary school.

Table 2.1 lists output and input variables used in the empirical analysis and Table 2.2 presents summary statistics for the main outcome variables - earnings and adult cognitive skills, individual background factors, early cognitive ability, early health status, and instrumental variables. In Table 2.2 I break down the mean values associated with schooling levels (Column 2 and Column 3) and present the results of testing for mean differences (Column 4). For all of the outcomes, there is a clear pattern by upper secondary school enrolment. Those attended upper secondary school have higher earnings and have higher cognitive ability at adulthood. Individual characteristics and background factors prior to upper secondary school entrance are also remarkably different between the two groups. The former have higher stocks of early cognitive ability, healthier, come from wealthier families, and live in communities with better infrastructure.

	S=0 not attend	S=1 attend	Mean difference $Y(0) - Y(1)$ (std. err.)
Outcome variables			
- Annual income ($/year, 2006 price)	873.393	1619.083	-745.689*** (62.568)
- Adult cognitive ability (latent)			
Observed covariates			
Early cognitive ability (latent)	0.925	1.134	-0.208*** (0.007)
Early health status (standardized)	-0.149	0.098	-0.248*** (0.024)
Family wealth index before age 13	0.517	0.678	-0.160*** (0.004)
Age	9.646	9.200	0.446*** (0.075)
Gender (male =1)	0.504	0.507	-0.003 (0.01)
Family size	6.812	6.609	0.203*** (0.058)
Community-level wealth index before age 12	0.474	0.541	-0.067*** (0.004)
Instrumental variables			
GPS distance from nearest upper secondary school to surveyed cluster (kms)	4.464	2.873	1.591*** (0.099)
Total number of upper secondary schools accessible to community residents	3.489	4.268	-0.778*** (0.054)
N	5209		

Table 2.2: Descriptive statistics: Main data

2.4 Empirical results

2.4.1 The determinants of schooling choices

I first predict the propensity score $\hat{P}(\boldsymbol{Z})$ from a probit model of upper secondary school attendance with $(\boldsymbol{X}, \Theta_1)$ and Z_+ as regressors. I use a flexible probit specification as reported in Table 2.3. Alternative specifications for the schooling choice equation, including logit or linear probability model of $P(\boldsymbol{Z})$, or with alternative Z_+ (excluding either the nearest distance or the number of accessible upper secondary schools), do not alter the results I discuss here.

Table 2.3 reports the coefficients from the first stage estimation. As expected, the nearest distance and the number of accessible schools are strong predictors of upper secondary schooling choice. The coefficients of the exclusion restrictions reveal a positive relationship between the supply of upper secondary schools and the decision to attend. The closer the nearest upper secondary school and/or the higher number of accessible upper secondary schools in the community, the more likely one goes to upper secondary school.

Turning to individual characteristics, children with higher stocks of early cognitive ability, early health, and from wealthier families are more likely to attend upper secondary school. Although the test scores in this study provide information on multiple aspects of cognitive skills (intellect, reasoning, numeracy and literacy), and cognitive ability is quantitatively important, controlling for them does not substitute for the role of family socioeconomic status (SES) as measured by family wealth. These results altogether imply a critical role for family SES in driving schooling decisions in Indonesia, regardless of student abilities.

Moreover, the coefficients of the interactions between early capabilities - early cognition and health status - and family wealth are positive and statistically significant, pointing to a strong complementarity between the two characteristics on individual schooling choices. Figure 2.2 illustrates the effects of the complementarities on schooling choices by plotting the contour plots of the propensity scores (i.e., the probability of selecting into upper secondary school) on two dimensions - capabilities (cognitive or health) and wealth.

I proceed by investigating the density of the propensity scores $\hat{P}(\boldsymbol{Z})$. The first-stage schooling choice model generates a large common support of $\hat{P}(\boldsymbol{Z})$ from from 0.09 to 0.97, allowing us to identify MTE as the unobserved resistance approaches zero or one. Figure 2.1 shows the unconditional support generated by variation in the instrumental variables Z_+ and the covariates $(\boldsymbol{X}, \Theta_1)$. Under Assumptions (1) and (3), the MTE is additively separable in $(\boldsymbol{X}, \Theta_1)$ and $\mathcal{V}$, and identified from the marginal support of $\hat{P}(\boldsymbol{Z})$ as opposed to the conditional support. The supports of the predicted propensity scores overlap almost everywhere, although they are scattered and thin at the two tails of the distributions. Following Carneiro et al. (2011) and Brinch et al. (2017), I trim the data by dropping 53 observations for which there is limited common support, which correspond to the 0.01 percentiles and 0.09 percentiles in the $\hat{P}(\boldsymbol{Z})$ distributions given $S = 1$ and $S = 0$, respectively.

2.4.2 The marginal returns to upper secondary school on the labour market

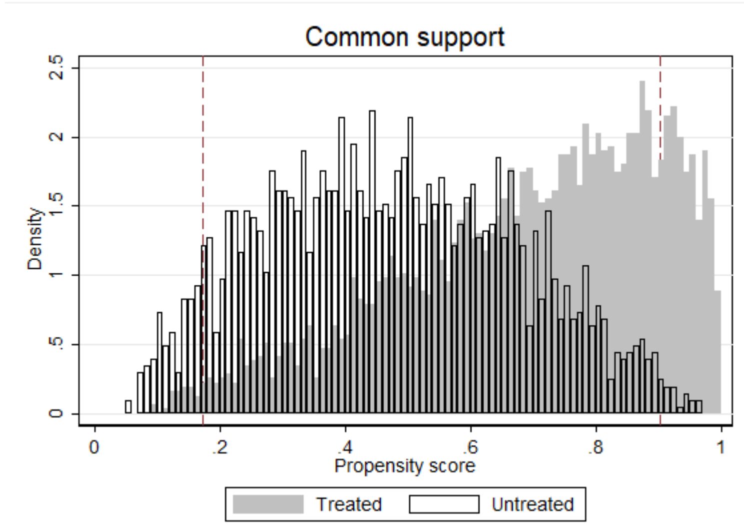
2.4.2.1 Testing for the presence of selection on gains

In Table 2.4 I present a series of tests for the presence of selection on unobserved gains for the earning outcome using the method developed in Heckman, Schmierer and Urzua (2010). The null hypothesis is whether the MTE is constant with respect to unobserved resistance $\mathcal{V}$, i.e., in Equation (2.10) $E(\triangle U|\mathcal{V} = K'(p)$ is flat. If $K'(p)$ is flat, i.e., it does not depend on propensity scores, the heterogeneity of marginal returns is not important, or students do not self select into upper secondary school attendance on the basis of unobserved earning gains. The test procedure is as follows. I specify $K(P)$ in Equation (2.9) as a polynomial of P of order κ and estimate the MTE using the local polynomial estimator, which is described in Appendix A.2.2. I then test whether the coefficients on the terms $\kappa + 1$ are jointly equal to zero. Rejecting the null hypothesis indicates that $K'(p)$ is not flat in p, and as can be seen from Equation (2.9), this suggests that $MTE(x, \theta, p)$ varies in terms of unobserved earning gains.

Upper secondary schooling choice equation

	Probit coef. (Std. err.)
Distance to nearest upper secondary school (log)	-0.055*** (0.016)
Total number of upper secondary schools (number)	0.023*** (0.009)
Early cognitive skills (latent)	0.492** (0.235)
Early health status (standardized)	0.127*** (0.065)
Family wealth index	1.197*** (0.391)
Early cognitive skills × family wealth	0.950*** (0.373)
Early health status × family wealth	0.334*** (0.107)
Gender (male = 1)	-0.036 (0.039)
Family size	-0.020*** (0.008)
Age	-0.049*** (0.007)
Commune-level infrastructure index	0.427*** (0.136)
Constant	-1.227*** (0.256)
N.	5209

Table 2.3: First-stage equation of schooling choices



s

Figure 2.1: Empirical support of propensity scores.

The graph plots the frequency distribution of $\hat{P}(\boldsymbol{Z})$ by schooling choices, $S = 0$ versus $S = 1$. The propensity score is estimated using a probit model with regressors in column 1 of Table 2.3. The support of $P(\boldsymbol{Z})$ ranges from 0.06 to 0.99. The common support, i.e., the overlapping region of $\hat{P}(\boldsymbol{Z})$ by schooling status, is from 0.09 to 0.97. The region between two vertical lines, from 0.18 to 0.89, is the common support on which I estimate the MTE and other treatment effects, resulting from trimming 1 percent of treated and untreated subsample.

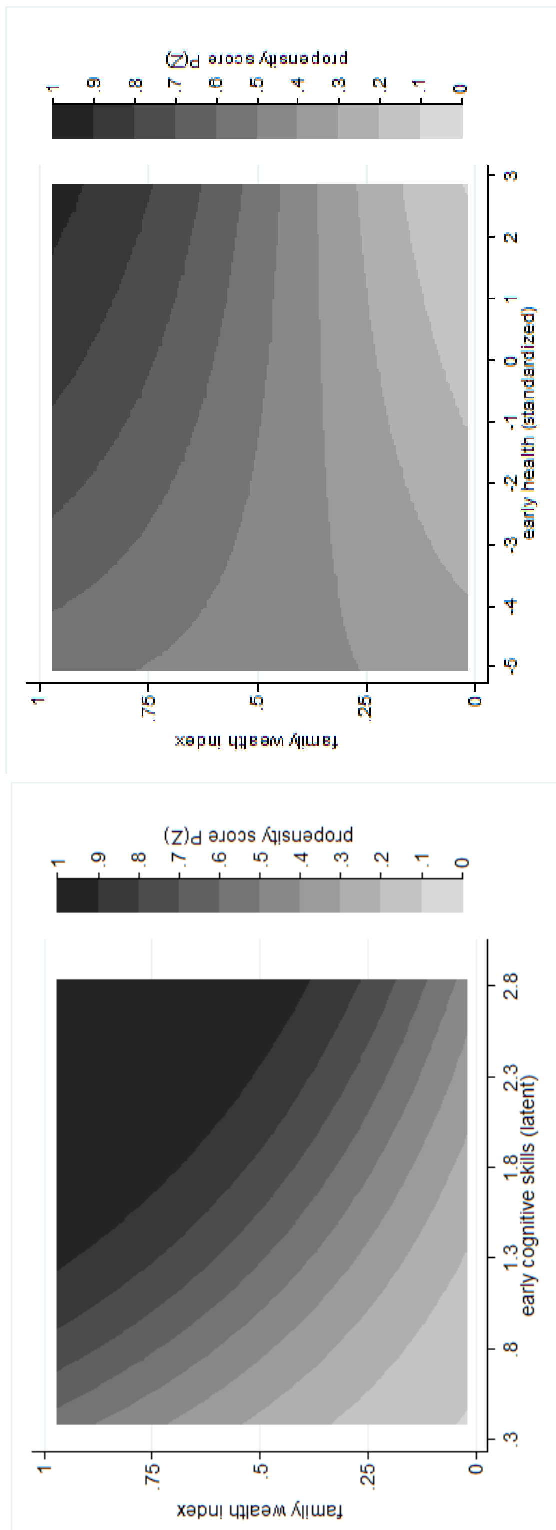
Polynomial order κ of $K(P)$	$\kappa = 2$	$\kappa = 3$	$\kappa = 4$	$\kappa = 5$
Joint test of coefficients of polynomials of $K(P)$ equal to zero ($\chi^2(\kappa)$)	3.30*	5.61*	10.03**	10.81**
$p-$value	0.0691	0.0606	0.0183	0.0287

Table 2.4: One-sided test for the presence of essential heterogeneity

In each column of Table 2.4, I specify $K(p)$ as a polynomial of orders $\kappa = 2, 3, 4, 5$ and present the $p-$values of joint tests that the coefficients on the terms $\kappa + 1$ are jointly equal zero. I account for the uncertainty in estimated propensity scores $\hat{P}(\boldsymbol{Z})$ by using the bootstrap. In all specifications, I use 250 bootstrap replications and in each iteration, I re-estimate the first stage $\hat{P}(\boldsymbol{Z}) = \Phi(\boldsymbol{Z}\gamma)$. The test results show that the null hypothesis of uncorrelated S and $\triangle U$ is rejected. That is, the marginal effects of attending upper secondary school on students with different degrees of unobserved resistance to schooling are heterogenous, which implies that students self-select into upper secondary school based partially on perceived *ex-post* pecuniary gains.

2.4.2.2 The marginal returns to upper secondary school

Figure 2.3a depicts a pattern of selection on pecuniary gains in terms of individual unobserved characteristics. The MTE curve relates the unobserved component of annual earnings, $\triangle U = U_1 - U_0$, to the quantiles $\mathcal{V}$ of the unobserved parts of upper secondary schooling choice. The MTEs are estimated from the full semiparametric model. The 90 percent confidence intervals are computed from a bootstrap with 250 replications. Higher values of $\mathcal{V}$ imply lower likelihood to attend upper secondary school, $\hat{P}(\boldsymbol{Z})$, and $\mathcal{V}$ represents the quantiles



(a)

(b)

Figure 2.2: Interactions between early capabilities and family wealth.
Sub-figure a: Interactions between early cognitive ability and family wealth. Sub-figure b: Interactions between early health status and family wealth.

of the distribution of individual unobserved "distastes" or "resistance" to upper secondary school. Figure 2.3a indicates that the marginal returns to upper secondary school decreases as the individual's distaste to schooling increases. Overall, regarding unobserved characterisics, students who are most likely to attend upper secondary school appear to benefit the most on the labour market in terms of annual earnings.

The extent of heterogeneity in earnings returns to upper secondary school is substantial: for the 30 percent of individuals who are least likely to attend upper secondary school, i.e., those located at the quantiles $\mathcal{V} > 0.7$, the marginal earnings returns to upper secondary school are negative albeit only marginally significant (see Figure 2.3a). By contrast, the earnings returns for 53 percent individuals with lower degree of resistance to high schol ($\mathcal{V} < 0.53$) are not only positive but also statistically significant). As an example, persons located at the top quantile of unobserved resistance, i.e., $\mathcal{V} = (0.89, 0.90)$, incur a loss of about 29% in annual earnings per upper secondary schooling year, whereas, those near the bottom quantile of $\mathcal{V}$, e.g., $\mathcal{V} = (0.30, 0.31)$, benefit substantially from upper secondary school with returns to a year schooling being about 44%.

Given the observation that the MTE curve is downward sloping (Figure 2.3a), it is informative to directly test two hypotheses: (i) whether the MTE is constant in $\triangle U$ using the estimated MTE, (ii) whether the MTE slope is negative in $\triangle U$, that is, whether there is selection on unobserved gains to upper secondary school. The two tests are complementary to the previous two-sided test of selection on gains based on specifying $K(P)$ in Equation 2.9 as a nonlinear function of P, which does not require estimating the MTE. To do so, I evaluate the MTE in 10 equally spaced intervals between 0.08 and 0.96 (the range of common support of $\hat{P}(\boldsymbol{Z})$). As in Carneiro et al. (2011) and Brinch et al. (2017), I construct pairs of adjacent intervals, and take the mean of the MTE within each interval. The values obtained, $\triangle LATE_{i,i+1}$, are also local average treatment effects (LATEs) at different quantiles $\mathcal{V}$ of the unobserved resistance. The difference of LATE in interval i and LATE in adjacent interval $i+1$ is

$$\begin{aligned} \triangle LATE_{i,i+1} = \quad & E\left(\triangle Y|\boldsymbol{X} = \bar{x}, \Theta_1 = \bar{\theta_1}, LB_i \leq \mathcal{V} \leq UB_i\right) \\ & -E\left(\triangle Y|\boldsymbol{X} = \bar{x}, \Theta_1 = \bar{\theta_1}, LB_{i+1} \leq \mathcal{V} \leq UB_{i+1}\right) \end{aligned}$$

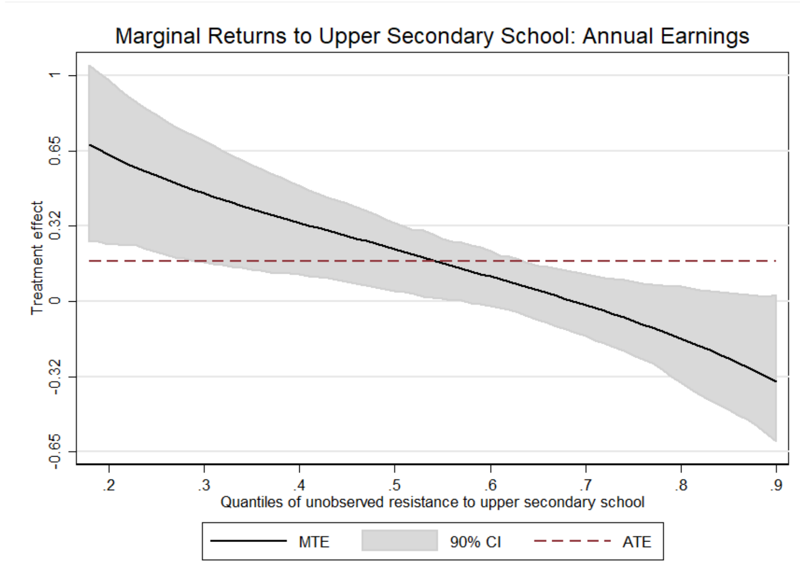
where LB and UB are the lower bound and upper bound of an interval, respectively. The first test corresponds to a two sided test, in which the null hypothesis is $\triangle LATE_{i,i'+1} = 0$. The second test is a one sided test in which the null hypothesis is $\triangle LATE_{i,i+1} \leq 0$ (MTE is non-decreasing) against the alternative that $\triangle LATE_{i,i'+1} > 0$ (MTE is downward sloping).

Table 2.5 shows that the null hypothesis of constant MTE (different LATEs over adjacent intervals) is rejected for all pairs at common levels of significance. Similarly, the last column of Table 2.5 indicates that the slope of the MTE cure is negative and statistically significant at common levels of significance for all values of $\hat{P}(\boldsymbol{Z})$ within the common support. This is the clearest evidence that individuals select into upper secondary school based on heterogenous returns in realized earnings, and the rejection of no selection on gains is strong in both the left and the right tails of the estimated MTE.

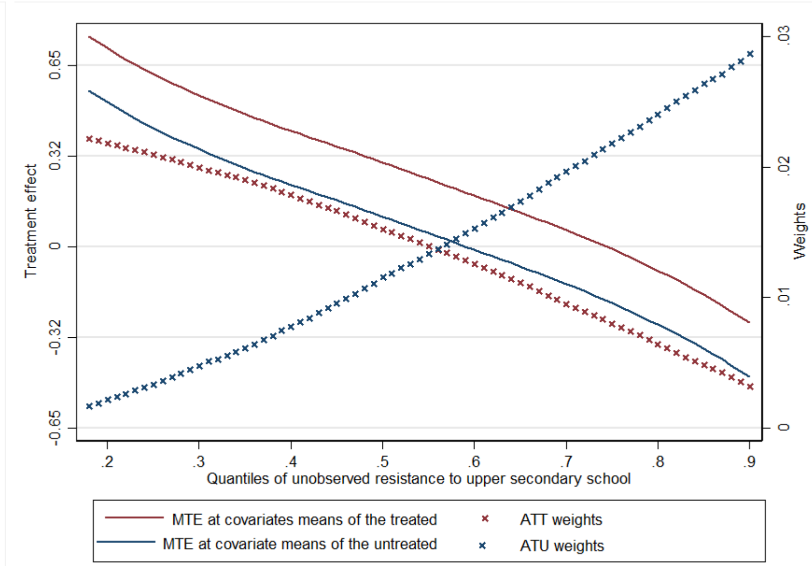
2.4.3 Summary measures of treatment effects and IV estimates

2.4.3.1 Summary measures of treatment effects

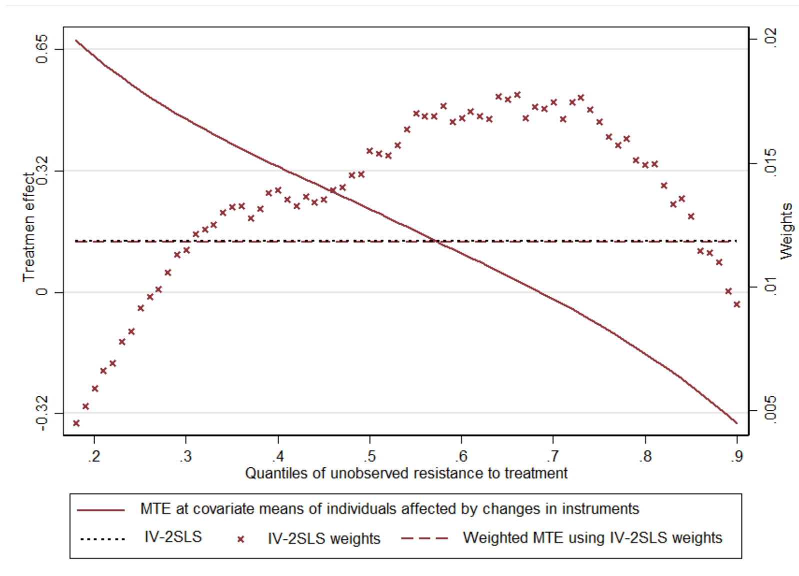
Column 1 of Table 2.6 presents the conventional causal effects of upper secondary school on annual earnings using the semiparametric LIV estimator. In principle, these average causal effects are obtained by appropriately aggregating over the MTE curve. Unfortunately, these population-level average parameters cannot be semiparametrically identified in my data, similarly to previous studies (Carneiro et al., 2011, 2015; Cornelissen et al., 2018), because the unconditional support of $\hat{P}(\boldsymbol{Z})$ isless than the unit interval. I follow Carneiro et al. (2011, 2015, 2017) to report the estimates of ATE, ATU, and ATT when the weights on $\hat{P}(\boldsymbol{Z})$ are restricted to integrate to 1 in the support of the MTE. The annualized returns to upper secondary schooling are produced by dividing these parameters by the difference in the average years of schooling of treated and untreated individuals, which equals to about 6.1 schooling years. I call these parameters $\tilde{ATE}$, $\tilde{ATT}$, and $\tilde{ATU}$ (to distinguish them from the original ATE, ATT and ATU which are not identified).



(a)



(b)



(c)

Figure 2.3: Annualized marginal returns to upper secondary school on annual earnings.
Estimated from full semiparametric MTE model. Sub-figure a: MTE curve evaluated at the weighted average of the covariates $(\boldsymbol{X}, \Theta_1)$. Sub–figure b: MTE curves and weights for the treated (ATT) and untreated (ATU) evaluated at the weighted averages of the covariates $(\boldsymbol{X}, \Theta_1)$. Sub-figure c: MTE curves and weights for individuals shifted schooling choice by the instruments.

LATE over intervals of $P(\boldsymbol{Z})$ supports	Point estimates (annualized) $\triangle LATE_{i,i+1}$ (std. err.)	p-value	
		Nonconstant MTE Ha: $\triangle LATE_{i,i+1} \neq 0$	Downward sloping MTE Ha: $\triangle LATE_{i,i+1} > 0$
(0.18, 0.21 - 0.22, 0.25)	0.076 (0.033)	0.022**	0.011***
(0.25, 0.28 - 0.29, 0.32)	0.058 (0.025)	0.018**	0.009***
(0.32, 0.35 - 0.36, 0.39)	0.050 (0.020)	0.011***	0.005***
(0.39, 0.42 - 0.43, 0.46)	0.044 (0.017)	0.010***	0.005***
(0.46, 0.49 - 0.50, 0.53)	0.046 (0.016)	0.004***	0.002***
(0.53, 0.56 - 0.57, 0.60)	0.048 (0.017)	0.006***	0.003***
(0.60, 0.63 - 0.64, 0.67)	0.048 (0.019)	0.010***	0.005***
(0.67, 0.70 - 0.71, 0.74)	0.052 (0.021)	0.014**	0.007***
(0.74, 0.77 - 0.78, 0.81)	0.062 (0.026)	0.020**	0.010***
(0.81, 0.83 - 0.85, 0.88)	0.073 (0.035)	0.039**	0.020**

Table 2.5: Testing for nonconstant and downward sloping MTE by comparing LATEs across $P(\boldsymbol{Z})$ intervals

Summary of average returns to upper secondary school on annual earnings (annualized)	
Annual earnings (log)	Semiparametric LIV estimator
$\tilde{ATE}$	0.172** (0.086)
$\tilde{ATT}$	0.362*** (0.125)
$\tilde{ATU}$	-0.107 (0.101)
N	5156

Table 2.6: Average causal effects of upper secondary schooling on earnings (annualized)

The annualized $\tilde{ATE}$ is equal to 0.172, computed as an equally weighted average over the MTE curve in Figure 2.3a and evaluated at mean values of $\boldsymbol{X}$ and Θ_1. This $\tilde{ATE}$ implies that for an individual picked at random from the population of IFLS respondents, each year of upper secondary school raises annual earnings by about 18 percent. The estimated $\tilde{ATE}$ is, indeed, significantly different from zero at the 5 percent level of significance.

To compute the $\tilde{ATT}$ and $\tilde{ATU}$, respectively, I aggregate over the MTE curves evaluated at the mean values of $\boldsymbol{X}$ and Θ_1 of the treated and untreated subgroups (see, Cornellissen et al., 2016; Andresen, 2018). Figure 2.3b clearly shows that, at any values of unobserved resistance $\mathcal{V}$ (or alternatively, propensity scores $\hat{P}(\boldsymbol{Z})$), the MTE curve of those not going to upper secondary school lies below the MTE curve for those attending, reflecting the patterns of selection on gains based on characteristics $\boldsymbol{X}$ and early abilities Θ_1. Figure 2.3b also plots the weights applied to the MTE curves to compute the average effects of $\tilde{ATT}$ and $\tilde{ATU}$, respectively. While the $\tilde{ATT}$ is computed with highest weights given to low values of $\mathcal{V}$ (because individuals with low resistance to school are more likely to attend), the $\tilde{ATU}$ is heavily weighted at high values of $\mathcal{V}$ (because individuals with high resistance to school are less likely to attend). The findings for the $\tilde{ATT}$ suggests that for the average treated student, each year of upper secondary school results in about 32 percent higher annual earnings. Similar to the $\tilde{ATE}$, the effect is significantly different from zero at the 5 percent level. In contrast, attending upper secondary school does not result in positive earnings returns. Indeed, those individuals are likely to incur a loss of about 10 percent in annual earnings for each schooling year, but the effects are not statistically different from zero.

2.4.3.2 IV-2SLS estimate of returns to upper secondary school

As Heckman and Vytlacil (1999, 2005, 2007) demonstrate, the IV-2SLS parameter can be represented as weighted averages over the MTE curve as discussed in Section 2.2.2.2. Figure 2.3c plots the MTE curve evaluated at the mean values of $\boldsymbol{X}$ and Θ_1 for individuals who change their schooling choices in response to changes in the instruments (the red line) and the weights applied to the unobserved component of earnings (the red x line)..The IV-2SLS weights applied to the MTE curve are summarized in the last row of Table 2.7 in the Appendix.

As can be seen from Figure 2.3c the IV-2SLS estimator gives the largest weight to individuals with intermediate to high resistance to upper secondary school attendance. When applying these weights to the MTE curve, I obtain a weighted effect of 0.134 (dashed horizontal line in Figure 2.3c), which is close to the linear IV effect of 0.135 (dotted horizontal line) which is obtained from the two-stage least squares (2SLS) estimation. The similarity of IV-2SLS estimates is reassuring and can be considered a specification check for the MTE model. However, the conventional IV-2SLS estimate not only masks considerable heterogeneity in the response to treatment but also is difficult to interpret, especially in a setting that uses multiple continuous instrumental variables as in this study.

2.4.3.3 Robustness checks

Model validation using alternative instruments So far, I have used both the nearest distance and the total number of upper secondary schools accessible by commune residents as the instrumental variables. I now use each instrumental variable alternatively to validate the MTE estimates using both of them. The idea is to exploit the result in Equation 2.9 that the MTE is invariant to different instruments excluded from the potential outcome equations. If the MTE curves do not change significantly with excluded instruments, it will reinforce faith in the validity of the instruments.

Figure 2.4a compares the MTE curve using nearest distance to upper secondary school instrument to the MTE using both the nearest distance and the total supply instruments. In each case, I estimate the MTE semiparametrically using the LIV estimator. In particular, I fix the specification of school choice equation but change the excluded instruments from the outcome equation. In the main model I present above, the total number of upper secondary schools and the nearest distance are both excluded from the earnings equation. In this section, I use only the nearest distance as the instrument and exclude the total supply variable in the outcome and choice equation. The MTE curves in the two cases display the same downward sloping pattern in terms of unobserved characteristics. Indeed, the point estimates at each quantile of unobservables are very close in magnitude. This finding reassures that the differences in the IV estimates of returns to upper secondary school arise because the MTE is weighted differently for each instrument rather than because the instruments are invalid.

An alternative validating exercise is to estimate the MTE using only the supply variable as excluded instrument while excluding the nearest distance in both the schooling choice and earnings equations. Figure 2.4b shows the MTE curve using only the nearest distance. I use the same semiparametric LIV estimator as described above and keep the schooling choice equation fixed. As before, the MTE curve is decreasing over the quantiles of unobserved resistance to school, strengthening the credibility of the MTE estimates reported in Figure 2.3a.

Alternative specifications of MTE curves The patterns of selection on pecuniary gains with respect to both observed and unobserved characteristics is robust to several alternaltive specifications. Notice that I have already estimated the MTE by a flexible semiparametric specification, allowing for non-monotone changes of the MTE curve with individual's distaste for upper secondary schooling. I present MTE curves under alternative specifications of the MTE using the local polynomial estimator described in Appendix A.1, which is essentially parametric but does not impose any distributional assumptions on the unobservables. Figure 2.4c depicts MTE curves based on specifications of $K(P)$ as polynomials of degree $\kappa = 2$ (see Equation (2.15) in Appendix A.1). These curves are monotonically decreasing with distastes to upper secondary school attendance, with their shapes generally resembling the semiparametric MTE curve in Figure 2.3.

2.4.4 Interpretation and learning outcomes

2.4.4.1 Counterfactual wage outcomes and the source of the wage returns heterogeneity

Given the finding that students with the lower resistance to upper secondary school benefit more from attending upper secondary school on the labour market, I now attempt to shed light on the pattern of selection on gains that the analysis implies. To summarize, I first investigate whether the decreasing earnings returns to upper secondary school by individuals' resistance to upper secondary school (that is, $E(U_1 - U_0|\mathcal{V} = v)$ in Equation (2.3) is driven by earnings returns in the untreated state (that is, $E(U_0|\mathcal{V} = v)$) or earnings returns in the treated state (that is, $E(U_1|\mathcal{V} = v)$). Specifically, I adopt the estimation procedure of Brinch et al. (2017) to semiparametrically estimate $E(U_0|\mathcal{V} = v)$ and $E(U_1|\mathcal{V} = v)$, which these quantities builds on the control function approach for the MTE model proposed by Heckcman and Vytlacil (2007) (the procedure is summarized in Appendix A.3).

Figure 2.5 presents the separate curves for the unobserved component of earnings returns in the treated and untreated state, in terms of resistance to schooling. The emerging pattern is remarkable: while the earnings

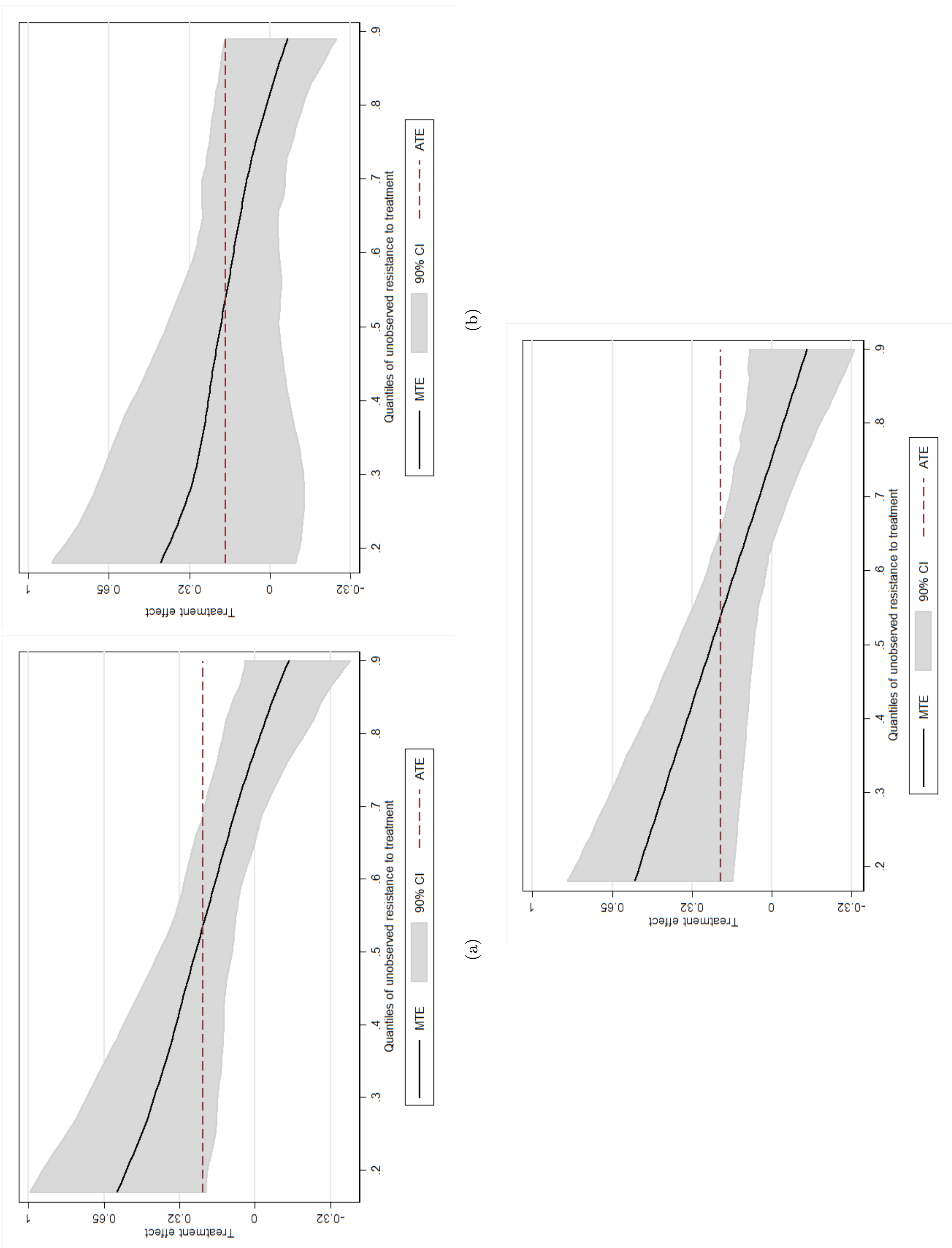


Figure 2.4: The MTE curves with different excluded instruments.
Sub-figure a: the MTE curve when only nearest distance is instrumental variable. Sub-figure b: the MTE curve when only total number of schools is instrumental variable. Sub-figure c: the MTE curve using the local polynomial estimator using two IVs simultaneously.

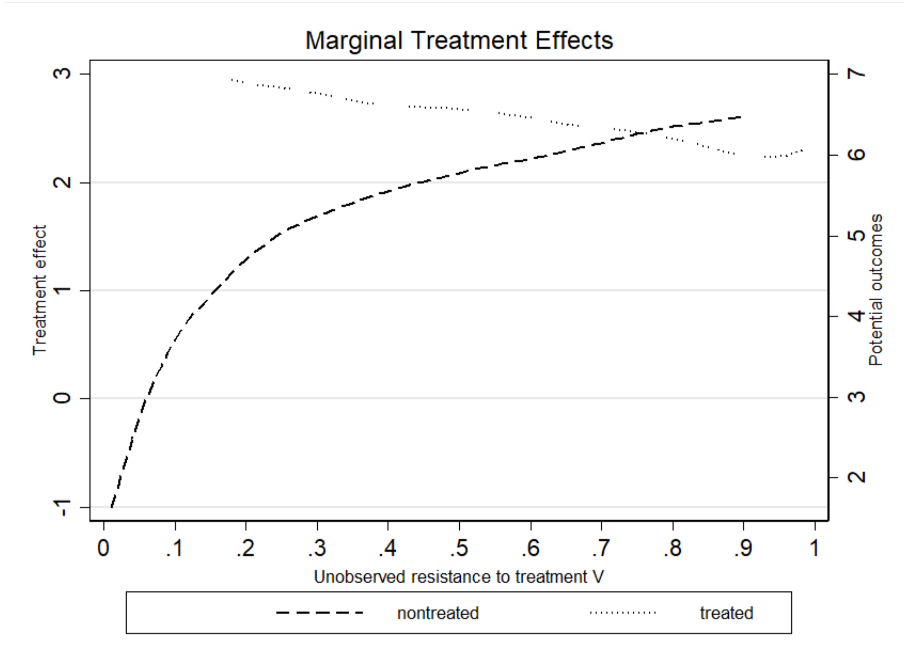


Figure 2.5: Counterfactual outcome (unobserved part) as a function of resistance to treatment, by treatment state.
The figure plots the unobserved component ofannual earnings against the quantiles $\mathcal{V}$ of unobserved resistance to treatment V, separately for the treated (i.e., $E(U_1|\mathcal{V} = v)$, dotted line) and untreated (i.e., $E(U_0|\mathcal{V} = v)$, dashed line) state, following Brinch et al. (2017).

returns to unobserved characteristics in the untreated state, $E(U_0|\mathcal{V} = v)$, is increasing everywhere, the returns in the treated state, $E(U_1|\mathcal{V} = v)$, is decreasing. This result suggests that higher marginal returns to upper secondary school of low-resistant students on the labour market are driven by their higher earnings returns in the treated state (decreasing $E(U_1|.)$ and lower earnings returns in the untreated state (increasing $E(U_0|.)$. The earnings returns for low-resistant students are significantly higher than their counterparts if attending upper secondary school, but lower without upper secondary school. Notice also that the decreasing $K_1(p)$ curve and the pattern of selection on observed characteristics altogether imply that upper secondary school acts as an *economic disequalizer* that perpeptuates the intergroup differences in earnings returns between those more likely to attend upper secondary school (low resistance, from wealthier families and having higher stocks of early capabilities) and those less likely to attend (higher resistance, from less wealthier families and having lower stocks of early capabilities).

Lastly, while the $E(U_1|\mathcal{V} = v)$ curve is downward sloping, it is noticeably flatter than the $E(U_0|\mathcal{V} = v)$ curve, leading one to ask why the highly resistant individuals earn less than the lowly resistant one when both of them attend upper secondary school. A common explanation for the latter result is that the high-resistance students may learn much less than those with low resistance, and thus, have lower stocks of cognitive abilities at adulthood, which will translate into economic inequalities on the labour market. In the next subsection, I formally test for this hypothesis.

2.4.4.2 The marginal returns to upper secondary school on learning outcomes

To investigate the question of whether the lower returns of high-resistance individuals in the labour market is driven by learning inequality, that is if they also learn much less from upper secondary school, I now assess the returns to schooling attendance on cognitive skills in adulthood. To this end, I estimate the MTE and average treatment effects using the LIV estimator and the sample of wage earners as before. The first stage estimation of propensity scores remains unchanged, as the location of individuals on the quantiles of unobservables $\mathcal{V}$ and their degrees of resistance to schooling attendance. The learning inequality would be present if I observe a pattern of selection on ability gains, which is similar to the pattern of selections on wage gains.

Remarkably, I find a completely reverse pattern of selection on abilities with respect to unobserved characteristics V as shown in Figure 2.6. Figure 2.6a provides evidence of reverse selection on gains in terms of unobserved characteristics. This figure shows the MTE curve for mean values of individual characteristics $\boldsymbol{X}$ and early cognitive skills Θ_1 in the main sample and relates the unobserved components of the treatment effect on adult cognitive skills and quantiles $\mathcal{V}$ of the unobserved component of schooling choices. The MTE curve increases with this resistance, completely contrary to the pattern of selection on wage gains found previously. Thus, on the basis of unobserved characteristics, individuals who are most likely to enroll in upper secondary schools appear to benefit the least from schooling attendance.

The reverse selection on unobserved cognition gains is reinforced by a similar reverse selection on observed cognition gains as shown in Figure 2.6b. The curve of marginal returns on cognitive skills, evaluated at the mean values of $\boldsymbol{X}$ and Θ_1 of those having attended in upper secondary school lies below the MTE curve at the $\boldsymbol{X}$ and Θ_1 of those who did not attend school. This reflects the reverse selection on observed cognitive gains. In summary, on the basis of observed characteristics $\boldsymbol{X}$ and early cognition Θ_1, individuals who are most likely to enroll in upper secondary schools appear to benefit the least from school attendance.

2.4.4.3 Interpreting the patterns of selections on pecuniary and nonpecuniary outcomes

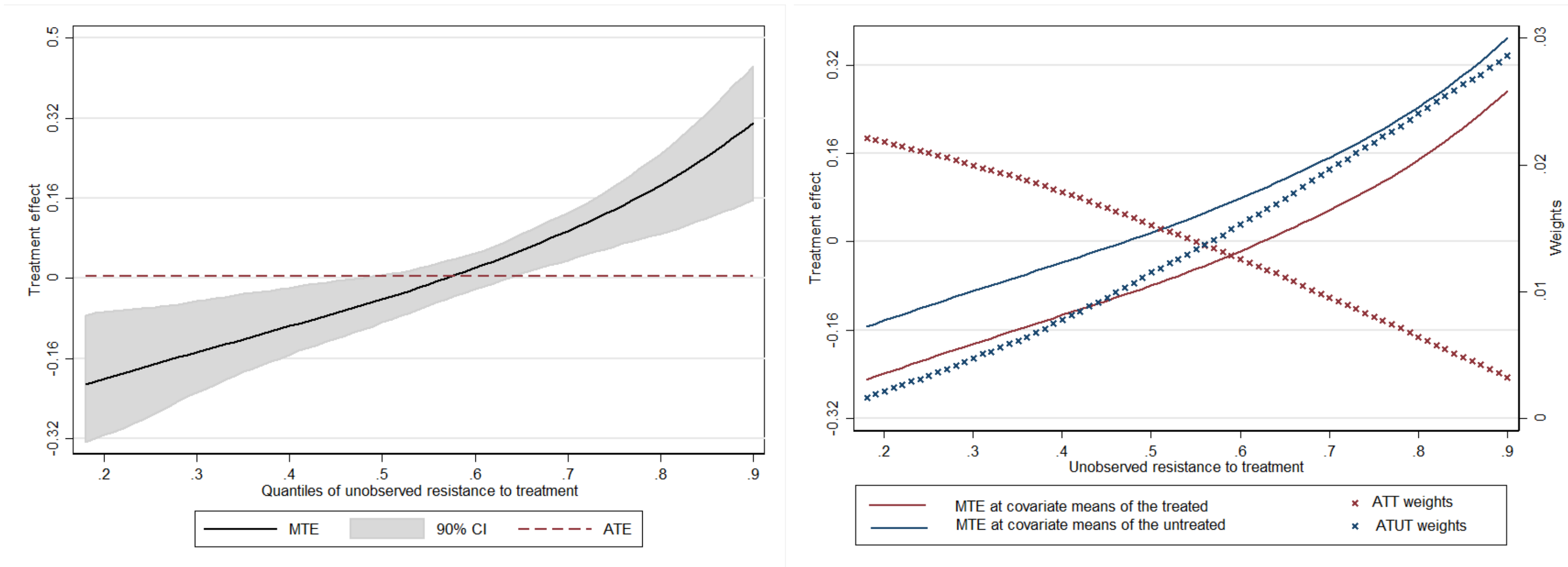
My findings on the patterns of selection on wage and learning outcomes give rise to an important question. Given the significant benefits to individual's cognitive abilities at adulthood, why disadvantaged individuals (high-resistance, coming from low SES backgrounds, having lower stocks of early health or early cognitive abilities) do not attend upper secondary schools more often. Or conversely why do better off individuals attend upper secondary schools even when there are no apparent benefits in terms of ability enhancement?

While individuals are likely to self-select into schools based on both pecuniary and nonpecuniary gains as suggested by previous studies (Attanasio et al., 2019; Beffy et al., 2012; Belfield et al., 2016; Boneva and Rauh, 2017, 2018), my analysis indicates that pecuniary gains are likely to be the driving force of the schooling choices by Indonesian youth. By contrast, the role of nonpecuniary gains, i.e., of cognitive development I identify in this paper is unclear. A possible explanation is that socio-economically disadvantaged students put higher weights on the pecuniary gains. Another cause for low rates of upper secondary school attendance by disadvantaged students could be that they are not informed about nonpecuniary gains, of which gains in cognitive ability are parts, as well as about pecuniary gains (wage gains). In addition, despite heavy subsidies from the Indonesian government, disadvantaged students may face higher costs of schooling relative to their family financial resources than advantaged students. This financial burden may further deter them from attending upper secondary school, because they cannot borrow against their future earnings.

2.5 Conclusion

In this paper, I assess the heterogeneity in the effects of upper secondary school on annual earnings and adult's cognitive capacity by estimating marginal returns to schooling attendance. Consistent with Carneiro et al. (2017), I find evidence that people select into upper secondary school on the basis of wage returns to schooling. In general, marginal and average returns to upper secondary school are not the same and conventional average return parameters. Building on a tighter identification strategy than usually adopted in Carneiro et al. (2017), I estimate the marginal returns to upper secondary school using a robust semi-parametric selection model. I document a substantial degree of heterogenous returns to upper secondary school attendance on pecuniary and nonpecuniary outcomes, with respect to both observed and unobserved individual characteristics.

For the main outcome I consider annual earnings, I find that children with unobserved characteristics that make them least likely to enter school benefit the least from schooling attendance in terms of wages. I then



(a) (b)

Figure 2.6: MTE curves for cognitive abilities at adulthood.

Sub-figure a: MTE curves for cognitive abilities at adulthood evaluated at mean values of individual characteristics and early cognitive skills of the sample. Sub-figure b: MTE curves for cognitive abilities at adulthood evaluated at mean values of individual characteristics and early cognitive skills of the treated and nontreated subsamples.

test for the importance of self-selection on wage gains on the labor market. The data suggest that self-selection on pecuniary gains is an empirically important phenomenon governing upper secondary schooling choices in Indonesia. Individuals sort into schooling on the basis of wage gains which are observed by the economist as well as unobserved (by the economist) variables. These results are robust regardless of the empirical specifications.

The findings on the marginal wage returns to upper secondary school raises the question of why students from advantageous backgrounds have higher marginal returns to high school degree and are more responsive to expansion of high schooling. I investigate whether the common explanation invoking the presence of learning inequality is valid one for the patterns of selection on pecuniary gains. I take advantage of cognitive measures at adulthood to study whether students from disadvantaged backgrounds and with higher resistance to school attendance also learn less than their better off counterparts. I show that there is little evidence for learning inequality in high school. The findings reveal that although early cognitive skills and advantaged family backgrounds (wealth) promote adult cognitive ability independently of schooling effects, attending high school not only promotes better cognition but also almost fully compensates for early deficiencies in those characteristics. In other words, students from disadvantaged backgrounds learn as much as those from advantageous backgrounds and learning inequality is unlikely the cause of the revealed inequality on the labour market.

Bibliography

[1] Aakvik, Arild; Heckman, James J.; Vytlacil, Edward J. Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs. Journal of Econometrics. 2005; 125(1–2):15–51.

[2] Attanasio, Orazio, Teodora Boneva, and Christopher Rauh. Parental Beliefs about Returns to Different Types of Investments in School Children. No. w25513. National Bureau of Economic Research, 2019.

[3] Beffy, Magali, Denis Fougere, and Arnaud Maurel. Choosing the field of study in postsecondary education: Do expected earnings matter?. Review of Economics and Statistics 94.1 (2012): 334-347.

[4] Belfield, C., Boneva, T., Rauh, C. and Shaw, J.. Money or fun? Why students want to pursue further education (2016).

[5] Boneva, Teodora, and Christopher Rauh. Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. (2017).

[6] Boneva, Teodora, and Christopher Rauh. Parental Beliefs about Returns to Educational Investments—The Later the Better?. Journal of the European Economic Association 16.6 (2018): 1669-1711.

[7] Björklund, Anders; Moffitt, Robert. The Estimation of earnings Gains and Welfare Gains in Self-Selection. Review of Economics and Statistics. 1987; 69(1):42–49.

[8] Cameron, Stephen V.; Taber, Christopher. Estimation of Educational Borrowing Constraints Using Returns to Schooling. Journal of Political Economy. 2004; 112(1):132–182

[9] Cameron, Stephen V.; Heckman, James J. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. Journal of Political Economy. 1998; 106(2): 262–333.

[10] Cameron, Stephen V.; Heckman, James J. The Dynamics of Educational Attainment for Black, Hispanic, and White Males. Journal of Political Economy. 2001; 109(3):455–99.

[11] Card, David. Using Geographic Variation in College Proximity to Estimate the Return to Schooling. National Bureau of Economic Research. 1993; 4483

[12] Card, David. Using Geographic Variation in College Proximity to Estimate the Return to Schooling. In: Christofides, Louis N.; Grant, E. Kenneth; Swidinsky, Robert, editors. Aspects of Labour Market Behaviour: Essays in Honor of John Vanderkamp. University of Toronto Press; Toronto: 1995. p. 201-222.

[13] Card, David. The Causal Effect of Education on Earnings. In: Ashenfelter, O.; Card, D., editors. Handbook of Labor Economics. Vol. Vol. 5. North-Holland; New York: 1999. p. 1801-1863.

[14] Card, David. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. Econometrica. 2001; 69(5):1127–1160.

[15] Carneiro, Pedro; Heckman, James J. The Evidence on Credit Constraints in Post-Secondary Schooling. Economic Journal. 2002; 112(482):705–734.

[16] Carneiro, Pedro; Heckman, James J.; Vytlacil, Edward J. Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. Econometrica. 2010; 78(1):377–394. [PubMed: 20209119]

[17] Currie, Janet; Moretti, Enrico. Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. Quarterly Journal of Economics. 2003; 118(4):1495– 1532.

[18] Fan, Jianqing; Gijbels, Irene. Local Polynomial Modelling and its Applications. Chapman and Hall; New York: 1996.

[19] Glewwe, P. . The relevance of standard estimates of rates of return to schooling for education policy: A critical assessment. Journal of Development economics, (1996); 51(2), 267-290.

[20] Glewwe, P. (2002). Schools and skills in developing countries: Education policies and socioeconomic outcomes. Journal of economic literature, 40(2), 436-482.

[21] Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. Handbook of the Economics of Education, 2, 945-1017.

[22] Hansen, Karsten T.; Heckman, James J.; Mullen, Kathleen J. The Effect of Schooling and Ability on Achievement Test Scores. Journal of Econometrics. 2004; 121(1-2):39–98.

[23] Heckman, James J. Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy. Journal of Economic Literature. 2010; 48(2):356–398. [PubMed: 21743749]

[24] Heckman, J. J., & Li, X. (2004). Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000. Pacific Economic Review, 9(3), 155-171.

[25] Heckman, James J.; Schmierer, Daniel. Tests of Hypotheses Arising In the Correlated Random Coefficient Model. Economic Modelling. 2010 Forthcoming.

[26] Heckman, James J.; Vytlacil, Edward J. Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. Proceedings of the National Academy of Sciences. 1999; 96(8):4730–4734.

[27] Heckman, James J.; Vytlacil, Edward J. The Relationship Between Treatment Parameters Within a Latent Variable Framework. Economics Letters. 2000; 66(1):33–39.

[28] Heckman, James J.; Vytlacil, Edward J. Local Instrumental Variables. In: Hsiao, Cheng; Morimune, Kimio; Powell, James L., editors. Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya. Cambridge University Press; New York: 2001a. p. 1-46.

[29] Heckman, James J.; Vytlacil, Edward J. Policy-Relevant Treatment Effects. American Economic Review. 2001b; 91(2):107–111.

[30] Heckman, James J.; Vytlacil, Edward J. Structural Equations, Treatment Effects and Econometric Policy Evaluation. Econometrica. 2005; 73(3):669–738.

[31] Heckman, James J.; Vytlacil, Edward J. Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation. In: Heckman, J.; Leamer, E., editors. Handbook of Econometrics. Vol. Vol. 6B. Elsevier; Amsterdam: 2007a. p. 4779-4874.

[32] Heckman, James J.; Vytlacil, Edward J. Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Economic Estimators to Evaluate Social Programs and to Forecast Their Effects in New Environments. In: Heckman, J.; Leamer, E., editors. Handbook of Econometrics. Vol. Vol. 6B. Elsevier; Amsterdam: 2007b. p. 4875-5144.

[33] Heckman, James J.; Schmierer, Daniel; Urzua, Sergio. Testing the Correlated Random Coefficient Model. Journal of Econometrics. 2010; 158(2):177–203. [PubMed: 21057649]

[34] Heckman, James J.; Ichimura, Hidehiko; Todd, Petra E. University of Chicago, Department of Economics; 1997. How Details Make a Difference: Semiparametric Estimation of the Partially Linear Regression Model. Unpublished manuscript

[35] Heckman, James J.; Ichimura, Hidehiko; Smith, Jeffrey; Todd, Petra E. Characterizing Selection Bias Using Experimental Data. Econometrica. 1998; 66(5):1017–1098.

[36] Heckman, James J.; Tobias, Justin L.; Vytlacil, Edward J. Four Parameters of Interest in the Evaluation of Social Programs. Southern Economic Journal. 2001; 68(2):210–223.

[37] Heckman, James J.; Urzua, Sergio; Vytlacil, Edward J. Understanding Instrumental Variables in Models with Essential Heterogeneity. Review of Economics and Statistics. 2006; 88(3):389–432.

[38] Heckman, James J.; Urzua, Sergio; Vytlacil, Edward J. Instrumental Variables in Models with Multiple Outcomes: The General Unordered Case. Les Annales d'Economie et de Statistique. 2008; 91-92:151–174

[39] Imbens, Guido W.; Angrist, Joshua D. Identification and Estimation of Local Average Treatment Effects. Econometrica. 1994; 62(2):467–475.

[40] Kane, Thomas J.; Rouse, Cecilia E. Labor-Market Returns to Two- and Four-Year College. American Economic Review. 1995; 85(3):600–614.

[41] Kling, Jeffrey R. Interpreting Instrumental Variables Estimates of the Returns to Schooling. Journal of Business and Economic Statistics. 2001; 19(3):358–364.

[42] Pettersson G. 2012. Do supply-side education programs targeted at under-served areas work? The impact of increased school supply on education and earnings of the poor and women in Indonesia. Working Paper No. 49-2012, Economics Department, University of Sussex.

[43] Psacharopoulos G, Patrinos H. 2004. Returns to investment in education: a further update. Education Economics 12(2): 111–134.

[44] Robinson P. 1988. Root-N-consistent semiparametric regression. Econometrica 56(4): 931–954. Schennach S. 2013. Measurement error in nonlinear models: a review. In Advances in Economics and Econometrics: Theory and Applications, Acemoglu D, Arellano M, Dekkel E (eds.). Cambridge University Press: Cambridge, UK; 296–337.

[45] Vytlacil E. 2002. Independence, monotonicity, and latent index models: an equivalence result. Econometrica 70 (1): 331–341.

[46] Wang X, Fleisher B, Li H, Li S. 2007. Access to higher education and inequality: the Chinese experiment. Working paper, IZA. Willis R, Rosen S. 1979. Education and self-selection. Journal of Political Economy 87(5): 27–36.

[47] Leckelt, Marius, et al. "Validation of the Narcissistic Admiration and Rivalry Questionnaire Short Scale (NARQ-S) in convenience and representative samples." Psychological assessment 30.1 (2018): 86.

[48] Palczyńska, Marta, and Karolina Świst. "Measurement Properties of Non-cognitive scales in the Polish Follow-up Study on PIAAC (POSTPIAAC)." (2016).

Appendix

A.2.1. Unconditional average treatment-effect parameters and weights

Parameter	Interval of quantiles $\mathcal{V}$ of the unobserved resistance V	Weights applied to covariates $(\boldsymbol{X}, \Theta_1)$	Weights applied to the unobserved gains
ATE	$(0, 1)$	1	$\frac{1}{\mathtt{d}}$
ATU	$\mathcal{V} \le p$	$\frac{p_i}{E(p)}$	$\frac{P(p>v)}{\mathtt{d}E(p)}$
ATT	$\mathcal{V} > p$	$\frac{1-p_i}{1-E(p)}$	$\frac{1-P(p>v)}{\mathtt{d}\{1-E(p)\}}$
LATE	$P(z_+) < \mathcal{V} \le P(\tilde{z}_+)$	$\frac{p(z_{+i})-p(\tilde{z}_{+i})}{E(\tilde{p})-E(p)}$	$\frac{P\{p(z_+)>v\}-P\{\tilde{z}_+)>v\}}{\mathtt{d}(\bar{\tilde{p}}-\bar{p})}$
IV-2SLS		$\frac{\{\tau_i-E(\tau)\}(S_i-\bar{S})}{cov(S,\tau)}$	$\frac{\{E(\tau\mid p>v)-E(\tau)\}P(p>v)}{\mathtt{d}\times cov(S,\tau)}$

Table 2.7: Unconditional average treatment-effect parameters and weights.
Note: The weights in Column 2 and 3 follow the derivation of Heckman and Vytlacil (2007) and Cornelissen et al. (2016). The distribution of $\mathcal{V}$ is discretized with $\mathtt{d}$ points. For the IV-2SLS parameter, τ_i measures how much the propensity scores for each individual i is affected by the instrumental variables. Individuals with large absolute values of τ are given higher weights. These individuals are also more likely to have their schooling status determined by the instrumental variables.

A.2.2 Estimation procedure

A.2.2.1 Semi-parametric LIV estimator

The second step estimates (β_j, α_j), $j = 0, 1$, in Equation (2.9), using a semi-parametric double residual regression procedure (Robinson, 1988). I start by fitting a set of nonparametric regressions of Y and each element of $\boldsymbol{X}, \Theta_1$, $\boldsymbol{X} \times \hat{P}$ and $\Theta_1 \times \hat{P}$ on $\hat{P}(\boldsymbol{Z})$, which produce a set of residuals $e_Y, e_{\boldsymbol{X}}, e_{X\times\hat{P}}, e_\Theta, e_{\Theta\times\hat{P}}$. Then, the estimation of (β_j, α_j), $j = 0, 1$, requires regressing the residualized outcomes Y on the residualized $\boldsymbol{X}$, $X \times \hat{P}$, Θ_1 and $\Theta_1 \times \hat{P}$, that is,

$$e_Y = e_{\boldsymbol{X}}\hat{\beta}_0 + e_{\boldsymbol{X}\times\hat{P}}(\widehat{\beta_1 - \beta_0}) + e_{\Theta_1}\hat{\alpha_0} + e_{\Theta_1\times\hat{P}}(\widehat{\alpha_1 - \alpha_0}) + \epsilon. \tag{2.13}$$

The third step of the LIV estimator involves estimating the derivative $k(v)$. Notice that Equation (2.9) can be rewritten as

$$K\left(\hat{P}(\boldsymbol{Z})\right) + \tilde{\nu} = \underbrace{Y - \left(\boldsymbol{X}\hat{\beta}_0 + \Theta_1\alpha_0 + \hat{P}(\boldsymbol{Z})\left[\boldsymbol{X}(\widehat{\beta_1 - \beta_0}) + \Theta_1(\widehat{\alpha_1 - \alpha_0})\right]\right)}_{\tilde{Y}},$$

assuming that $E\left(\tilde{\nu}|\hat{P}(\boldsymbol{Z}),\boldsymbol{X},\Theta_1\right)=0$. This equation implies that $K\left(\hat{P}(\boldsymbol{Z})\right)$ can be interpreted as $E\left(\tilde{Y}|P(\boldsymbol{Z})=\hat{P}(\boldsymbol{Z})\right)$ and suggests a two-step procedure to estimate the $k(\hat{p})$: (i) constructing the residual $\tilde{Y}$, (ii) regressing non-parametrically $\tilde{Y}$ on $P(\boldsymbol{Z})$ to produce $K(p)$[12]. Finally, the MTE as a LIV estimator is computed as below:

$$\widehat{MTE}(x,\theta,v)=x(\widehat{\beta_1-\beta_0})+\theta(\widehat{\alpha_1-\alpha_0})+\underbrace{\frac{\delta\widehat{K(p)}}{\delta p}|_{p=v}}_{k(v)} \tag{2.14}$$

A.2.2.2 Parametric polynomial estimator

While the semi-parametric LIV estimators are robust to parametric assumptions, they are often estimated with low precision. Therefore, in the empirical application, I also present the results from the MTE model imposing a parametric assumption on (U_0,U_1,V). I implement this approach by first estimating the treatment selection equation in as a probit model as before to obtain estimated propensity scores $\hat{p}$. The second step of this estimator is to model $K(P(\boldsymbol{Z}))$ as a polynomial in $P(\boldsymbol{Z})$ of degree κ and estimating the following outcome equation:

$$Y=\boldsymbol{X}\beta_0+\Theta\alpha_0+\hat{p}\left[\boldsymbol{X}\left(\beta_1-\beta_0\right)+\Theta\left(\alpha_1-\alpha_0\right)\right]+\sum_{\kappa=2}^{K}\tau_\kappa\hat{p}^\kappa+\varepsilon. \tag{2.15}$$

The MTE curve is then the derivative of Equation (2.15) with respect to $\hat{p}$. I assume a second-order polynomial in $\hat{p}$ in the specifications presented in the main paper but generally find similar results for $\kappa=3,4,5$. To assess whether treatment effects vary with the unobserved resistance to treatment, I run tests for the joint significance of the second- and higher-order terms of the polynomial (i.e., the τ_κ in Equation (2.9)).

[12]I use the notation $K(p)$ instead of $K(\hat{P}(\boldsymbol{Z}))$ to reflect the fact that $\tilde{Y}$ can only be regressed on a subset values of $\hat{P}(\boldsymbol{Z})$ where the common support of $P(\boldsymbol{Z})$ exists.

Chapter 3

The effects of birth order on child's capabilities development: origins and mechanisms

Abstract

We investigate the origins and mechanisms of birth order effects on cognitive skills, socio-emotional skills and health in Vietnam. Using a sample of children from the Young Lives study, we find strong evidence of negative birth order effects on parental investments and child capabilities, emerging very early in life. Parents invest significantly less money in second-born children from age one until age eight. Second-born children are less healthy, have lower stocks of cognitive skills and socio-emotional skills than the first borns. We go beyond the previous studies and decompose the long-term effects of birth order on child capabilities, showing the existence of three channels: (i) the production technology efficiency, (ii) the parental monetary investments, (iii) the self-productiveness of child capabilities. While the first mechanism is significant only in health production technology during adolescence, the contributions of the second and third mechanisms are sizeable for the three analysed capabilities at different stages of development.

3.1 Introduction

Does a child's birth order have a long–term impact on that child's capabilities development? The question of birth order effects has been investigated for more than 100 years. As early as 1875, by analysing a sample of British scientists, Francis Galton found that first-born sons were overrepresented in the Royal Society (Galton, 1875). He hypothesized that the earlier borns would receive more financial resources for education, be more likely treated as companions by parents and undertake more responsibility, and receive more parental attention. These preferential treatments would allow them to be intellectually successful. Fifty years later, Alfred Adler extended the birth order issue to personality development (Adler, 1928).

In recent years, a large number of empirical studies has documented that child cognitive abilities and labour-market outcomes decline with birth order (see, e.g., Behrman and Taubman 1986; Kessler 1991; Hanushek 1992; Kantarevic and Mechoulan 2006; Black et al., 2005, 2007, 2018; Pavan, 2016; Lehmann et al., 2018). First-borns are also likely to be physically healthier, although the effects are mainly documented for adults (e.g., Lundborg et al., 2014; Barclay and Kolk, 2013; Modin, 2002; Jelenkovic et al., 2013; Black et al., 2014; Howe et al., 2014). In contrary, the findings on the relationship between birth order and children's noncognitive skills are much less conclusive (see the reviews by Ernst and Angst, 1983; Schooler, 1972; Black et al., 2018; Lehmann et al., 2018), primarily due to the difficulty in measuring individual's noncognitive skills in large surveys.

Despite the increasing interest on the topic shown by the recent theoretical and empirical research, the origin and the mechanisms generating the birth order gaps in child and adult outcomes are far from understood.

On the one side, theories do not provide clear guidance on the mechanisms underlying birth order effects. Parents face different financial and informational contraints over the life course which may result in systematic differences in investment decisions in their offspring. For example, later-born children may receive higher investments if parental wages grow significantly over the life cycle (Parish and Willis, 1993). On the other side, these disparities may benefit earlier borns because they share family resources with fewer siblings, especially parental time resources (Zajonc, 1976; Zajonc & Marcus, 1975; Birdsall, 1991; Price, 2008). Additionally, birth order may affect efficiency of the capability formation process. For example, parents may acquire and accumulate better knowledge on child rearing over time which makes the technology of capability production more efficient for the later borns. Children of different birth order may also follow different strategies to compete for parental resources and attention (Howe et al., 2002; Phinney, 1986; Sulloway, 1996). As a consequence, the peer environment is likely to be different for earlier and later borns which, in turns, affects the development process of their capabilities.

On the other side, very few empirical studies have attempted to identify the sources of the observed differences in individual outcomes by birth order. Among those Price (2008), Monfardini and See (2016), Black et al. (2018), Lehmann et al. (2018), and Pavan (2016). Price (2008) and Monfardini and See (2016) document a negative relationship between higher birth order and parental quality time investments in the U.S setting. However, they both find that the differences in children's cognitive skills by birth order cannot be explained by differences in parental time inputs. In the same U.S. setting, Lehmann et al. (2018) and Pavan (2016) find that differential parental monetary investments across siblings account for a substantial part of the negative birth order effects in cognitive abilities. Black et al. (2007) demonstrate that the negative birth order effects on IQ test scores of Norwegian adults cannot be explained by the biological origin which implies disadvantageous endowments of later born children. Recently, Black et al. (2018) find similar results regarding the insignificant role of nature in explaining the negative birth order effects on personality traits of Swedish adults. They further document systematic differences in parental monetary investment behavior that parents supervise less and are less strict towards later borns.

Most of other aforementioned research contributions attempt only to identify the sources of birth order effects, but do not quantify their contributions. Pavan (2016) is the only study which identifies and quantifies role of different sources using a dynamic factor model, limiting the attention to child cognitive development. However, the author focuses only on the gaps induced by differences in parental monetary investments and speculates that the remaining gaps in cognitive skills might be explained by modelling the development of child cognitive and noncognitive skills jointly. Moreover, existing studies usually investigate the birth order effect on a single aspect of individual abilities in isolation, ignoring the interdependence among capabilities formation.The only exception can be found in Lehmann et al. (2018), looking at both cognitive and non-cognitive development, but resorting to a simpler static fixed effects model ignoring the issue of skills measurement.

This paper adopts a dynamic factor model approach to investigate birth order effects on multiple aspects of human capital development and quantifies the contributions of distinct mechanisms. Specifically, we estimate birth order effects on child physical health, cognitive skills and socio-behavioural skills over 12 years of childhood. Building on Cunha and Heckman (2008), Cunha et al. (2010), and Heckman and Mosso (2014), we estimate a multi-stage production model of child capabilities jointly with with parental monetary investment equations. In this model, existing stocks of capabilities, parental monetary investments, and family background factors are inputs that produce, together with birth order, future capabilities. We account and correct for endogeneity of investments exploiting the plausibly exogenous variation in investments generated by eligibility to receive welfare benefits. Within this framework, children's birth order may affect the human capital development through its effects on parental monetary investments, on initial capabilities, and on efficiency of capability production technology. The birth order effect can also accumulate over time if child capabilities are productive in promoting themselves.

Our data are drawn from four waves of the Young Lives study, which was conducted in 2000, 2006, 2009, and

2013 in Vietnam. The study provides information on child birth order, multiple observed measures for each dimension of child capabilities, and child-specific parental monetary investments, and family backgrounds. Furthermore, we use information on family living conditions available in every wave which allows us to construct simulated eligibility to enrol in national welfare program which serves as instrumental variables for parental monetary investments.

Within this comprehensive framework, this paper brings two novel contributions to the birth order literature estimating the birth order effect on child-specific parental monetary investments, on the one side, and quantifying the contribution of three possible mechanisms generating the birth order gaps, on the other side. Specifically, the three identified channels are: (i) the direct effect of birth order on the efficiency of the capability production function, (ii) the indirect effect of birth order on capabilities through its effects on parental monetary investments; (iii) the accumulation of birth order effects over time due to self-productivity of child capabilities.

Using data from the Young Lives study in Vietnam, this paper first establishes the presence of negative birth order effects on different measures of child health, cognitive ability, socio-behavioural skills, and on parental monetary investments. We document that firstborns are healthier and perform better on cognitive skills and socio-behavioural tests than later borns, although the patterns over time are different across capabilities. While the negative birth order effect on child health is perpetuated, the gaps in cognitive skills decline substantially over the 12 years of childhood. Investigating more deeply into the origins of the patterns, we find that the negative relationship between birth order and capabilities for Vietnamese children is driven mainly by differences in parental monetary investments until age 8, and by the accumulation of birth order effects through self-productivity of child capabilities. However, there is little evidence for the effect of birth order on initial capabilities before age 1 and on the efficiency of capability production function. The joint modelling of the development of child health, cognitive skills and socio-behavioral skills allows us to shed light for the first time on the role of the self-productivity channel in shaping the birth order gaps.

The paper is structured as follows. Section 3.2 describes the Young Lives study in Vietnam and our sample. Section 3.3 presents the production technology of capabilities, the decomposition of birth order effects, the identification strategy and the estimation approach. Section 3.4 discusses our main results, including the estimated parameters of the production function, the gaps in capabilities by birth order implied by the model, and the decomposition of the gaps into different components. Section 3.5 concludes.

3.2 Description of data: Young Lives study in Vietnam

We use the sample of younger cohort from the Young Lives study for Vietnam (YLVN). The study tracks around 2000 children from 2001 to 2013 and provides repeated measures of child's health, cognitive skills and socio-emotional skills. The entire YLVN sample in the starting wave consisted of 2000 children aged from zero to one and residing in 2000 households. Follow-up studies have been conducted in 2006, 2009 and 2013.

Of the original 2000 children, we exclude children with missing family-size information (86), missing all capability measures in four waves (94). The remaining sample consists of 1728 children from 1728 families. Of these children, 218 come from one-child families, 1056 from two-child families, and 454 from families with three children or more. Our main analysis is performed on the sample of two-child families, a choice motivated by the well documented heterogeneity of birth order effects by family size. Among them we further discard 95 children who have missing information about parental monetary investments and family backgrounds. The final sample includes 961 children of which 502 are firstborns and 459 are secondborns.

Apart from birth order and repeated measures of child capabilities, we observe child-specific parental monetary investments, parental characteristics, family wealth, and details about family material living conditions. Moreover, the YLVN oversamples poor families in rural areas and is representative for ethnic diversity, therefore providing a large sample of families eligible for welfare benefits. These critical features allow us to simulate family-specific eligibility for welfare benefits on which we base our identification strategy.

Measures of child capabilities are available from 2001 to 2013 at four points in time, while parents are asked how much they spend on YL children on a yearly basis (for the previous twelve months) starting from the second round in 2006. Very importantly for our purposes, the measure of parental monetary investment is child-specific and is given by the sum of three main categories: education-, health- and entertainment-related spending. Parents provide information on how much they have spent on YL children on education and healthcare. For the other categories, such as gifts and clothing, they are asked about the total spending on all children in the family and the corresponding shares of the YL child. We deflate the monetary values to the base year of 2006 and adjust for rural/urban CPI. Detailed longitudinal information on family demographics and parental education are available on the same basis from 2001 to 2013.

For child capabilities, we focus on: physical health, measured by anthropometric measures and caregiver's assessment; cognitive skills, measured by math, reasoning and verbal tests; and socio-emotional skills, measured by psychological tests of self-efficacy, self-esteem, positiveness, aspiration and relationship with parents. The health assessment is conducted in all four study waves, while cognitive skills are measured starting in 2006, and socio-emotional skills starting in 2009. The health indicators are highly comparable and similar across waves. Anthropometric z-scores are age-in-days adjusted and normalized to have mean of zero and standard deviation of one based on a random sample of healthy U.S children. For cognitive skills, we use the test scores generated from item response models (IRT) and standardize the scores within the YL sample. For socio-emotional skills, YLVN provides information on children's response to psychological questions across different domains, which we use to generate the IRT scores. Similar to cognitive measures, we also standardize these scores within the YL sample.

For family background factors, we consider parental education, a wealth index, household size, living location (residence) and child birth order. The wealth index is a continuous measure between 0 and 1, constructed by the Young Lives team for every study wave. The index consists of three equally-weighted sub-indices: (i) housing quality, (ii) access to services and (iii) consumer durables. The wealth index captures the long-run material well-being of families and can also be interpreted as a measure of the living environment or the public goods that the child consumes with his/her family. The YLVN also includes information on housing materials which we use, together with the wealth index and living location, to simulate family eligibility to participate in welfare programmes, obtaining instrumental variables to address the endogeneity of parental monetary investments.

Table 3.1 presents the sample statistics of all variables used in the main analysis by birth order. To simplify the descriptive statistics, we compare the values of these variable averaged across all stages of child development. The summary statistics provide some interesting insights. While there is no clear indication for differences in health measures by birth order, the measures of cognitive skills and socio-behavioral skills are in general lower for second-born children. There is little evidence for differences in parental monetary investments and family background factors such as family wealth, caregiver's education.

3.3 The model

3.3.1 The empirical model

Our empirical model consists of three parts. The first one is a dynamic factor model in which child health, cognitive skills, and socio-emotional skills in the future are produced by existing child capabilities, parental monetary investments, child birth order, and family backgrounds. The second part approximates the parental monetary investment decision as a function of child capabilities, child birth order, family characteristics, and eligibility for welfare benefits. Welfare eligibility plays a crucial role in this model, allowing us to correct for endogeneity of parental monetary investments in the capability production technologies. The third part of the model is a measurement model we use to extract information about child unobserved capabilities from observed measures. Our empirical analysis of production technology and identification strategy largely builds on Cunha et al. (2010) and the recent work by Attanasio et al. (2017a, 2017b, 2018), and Agostinelli and Wiswall (2016a, 2016b).

Central to our model is the dynamic production technology of child capabilities. We assume Cobb-Douglas technologies of the following form:

Table 3.1: Main sample descriptive statistics

Variable	Total sample Mean (std. dev.)	First-born Mean (std. dev.)	Second-born Mean (std. dev.)
Child characteristics			
Parental monetary investments	665,196 (733,656)	692,661 (755,742)	635,158 (708,316)
(averaged over 12 years, in VND)			
Family wealth (index)	0.614 (0.146)	0.609 (0.148)	0.619 (0.144)
Caregiver's education			
High school	0.468 (0.499)	0.466 (0.499)	0.471 (0.500)
University	0.209 (0.407)	0.211 (0.409)	0.207 (0.406)
Child age			
HSP eligibility (simulated scores)			
Health measures			
height-for-age z-score	-0.980 (0.900)	-0.981 (0 .902)	-0.980 (0.898)
bmi-for-age z-score	-0.478 (0.944)	-0.431 (0.926)	-0.529 (0.962)
carergiver's assessment (1-3: bad-good)	2.115 (0.445)	2.130 (0.441)	2.099 (0.451)
Birth weight (standardized)	0.035 (0.993)	-0.154 (0.973)	0.248 (0.972)
Cognitive skills			
PPVT score (standardized)	0.129 (0.704)	0.156 (0.726)	0. 099 (0.679)
CDA score (standardized)	0.037 (0.998)	0.072 (0.975)	0.069 (0.863)
PIAT-Math score (standardized)	0.104 (0.825)	0.136 (0.789)	-0.001 (1.022)
EGRA score (standardized)	0.065 (0.969)	0.138 (0.967)	-0.015 (0.965)
Reading score (standardized)	0.114 (0.959)	0.154 (0.927)	0.070 (0.993)
Sociobehavioral skills			
Self-efficacy score (standardized)	0.058 (0.707)	0.107 (0.702)	0.004 (0.709)
Self-esteem score (standardized)	0.022 (0.697)	0.028 (0.682)	0.015 (0.715)
Positivity score (standardized)	0.062 (0.769)	0.067 (0.760)	0.056 (0.779)
Aspiration score (standardized)	0.057 (0.660)	0.110 (0.567)	-0.002 (0.745)
Child-parental closeness score (standardized)	0.026 (0.989)	0.024 (1.006)	0.028 (0.971)
Observations	961	502	459

$$\begin{aligned}\theta_{k,s+1} &= T_{k,s}\left(\theta_{H,s}^{\gamma_{k,H,s}}\theta_{C,s}^{\gamma_{k,C,s}}\theta_{N,s}^{\gamma_{k,N,s}}I_s^{\gamma_{k,I,s}}\right) \\ T_{k,s} &= \exp\left(\gamma_{k,0,s}+\gamma_{k,b,s}BO+\gamma_{k,X,s}X_s+\epsilon_{k,s}\right) \quad \text{with } T_{k,s}>0\end{aligned} \tag{3.1}$$

in which $k \in \{H, C, N\}$, denoting Health, Cognitive and Non-Cognitive skills respectively, $s = 1, 2, ..., S$, indicates the developmental stage, BO is a dummy variable with $BO = 1$ if the child is second born and $BO = 0$ otherwise; $T_{k,s}$ is a dynamic TFP term, which represents the technological level at stage s and depends on child birth order BO, family wealth w_s, background factors X_s, and unobserved idiosyncratic shocks $\epsilon_{k,s}$. In the empirical analysis we model the development of child capabilities over four stages from age 1 to age 12.

Rewritten in a log-linear fashion Equation (3.1) becomes:

$$ln\theta_{k,s+1} = \gamma_{k,0,s} + \gamma_{k,b,s}BO + \sum_{\ell}\gamma_{k,\ell,s}ln\theta_{\ell,s} + \gamma_{k,I,s}lnI_s + \gamma_{k,X,s}X_s + \epsilon_{k,s} \tag{3.2}$$

with $k, \ell \in \{H, C, N\}$. Notice that the parameters of the production technology are allowed to be stage and capability-specific. In each stage s, parents make investments I_s, which, together with existing stocks of child capabilities $\theta_{k,s}$ and other inputs, produce child capabilities in $s+1$. Given our primary aim of disentangling the impact of birth order on child capabilities, we have to account for two sources of estimate biases: the endogeneity of parental monetary investments and the unobservability of capabilities. We describe below our empirical strategy to deal with the first issue and postpone the discussion of the second issue to Section 3.3.3.1.

We estimate the production technology (3.2) jointly with an approximation of the investment decision. Specifically, we augment the block of three equations in (3.2) with the following investment equation:

$$lnI_s = \gamma_{I,0,s} + \gamma_{I,b,s}BO + \sum_{\ell}\gamma_{I,\ell,s}ln\theta_{\ell,s} + \gamma_{I,X,s}X_s + \gamma_{I,HSP,s}HSP_s + \epsilon_{I,s} \tag{3.3}$$

in which HSP_s is a family-specific measure of eligibility to welfare benefits at stage s. The idea is to approximate parental monetary investment as a function of the child capabilities, family background factors, and to use HSP eligibility as stage-specific instrumental variable for endogenous investments. Endogeneity of investments amounts to non-zero correlation of the investment error components $\epsilon_{I,s}$ with each of the $\epsilon_{k,s}$ for all $k \in \{H, C, N\}$. Similar to the parameters of the production technologies, we allow parameters in Equation (3.3) to be stage-specific.

In Equation (3.3) the variable HSP_s serves as a stage-specific instrumental variable which, conditional on (θ_s, w_s, X_s), generates exogenous variations in parental monetary investments and allows for consistent estimation of $\gamma_{k,I,s}$. Following Bernal and Keane (2011), we construct simulated eligibility for HSP participation instead of using actual HSP participation[1] or actual HSP eligibility[2]. I expect eligible parents, who are unable to insure against their own income and likely make suboptimal investment decisions would change their investments under the prospect of being eligible for HSP benefits. The identification of investment coefficients in the production technology (3.2) is then straightforward, using standard arguments for instrumental variables with HSP_s serving as stage-specific instruments for I_s in each stage s.

Joint estimation of Equation (3.2) and (3.3) is crucial to characterize the origins of birth order effects on parental investments, and to disentangle the mechanisms generating the effects on investments and child capabilities. Indeed, the approximation of the investment with birth order as a explanatory variable allows us to characterize the changes of birth order effects on investments across different development stages. Notice

[1] In principle, the actual benefits received (HSP participation) shift the household budget upwards due to the unconditional cash transfers, which must tie to improving living conditions and so family wealth. Consequently, this may induce two distinct effects: (i) it may relax liquidity constraints for the constrained families; (ii) the improvement in living conditions may have direct impact on child capabilities, especially child health. However, as we consider simulated eligibility, the latter possibility can be safely sidestepped.

[2] A downside of using the actual HSP eligibility to study parental monetary investment responses is that propitious selection or manipulation into eligibility into welfare programmes by families or resource misallocation by authorities, commonly documented for welfare reception, likely bias the estimates of α_s (McCrary, 2008; Miller et al., 2013).

that parental investments have been recognized as likely affected by birth order, but very few studies have estimated the evolution of birth order effects on these investments. Moreover, since the birth order gaps on parental investments and child capabilities might be generated through multiple channels, joint estimation of Equation (3.2) and (3.3) allow us to disentangle these effects and quantify their contributions. For example, if monetary investments are affected by birth order and we do not model them jointly but just include them as regressor in Equation (3.2), the coefficient $\gamma_{k,b,s}$ is only informative about the birth order effect on production efficiency and it is not possible to quantify the contribution to the birth order gaps on capabilities which occurs through differential investments.

The production technology we specify builds on Cunha and Heckman (2008) for process of skills formation, and to Pavan (2016) for birth order effects and inclusion of the parental investment decision in the model. However, we depart from both studies in two important directions. First, we extend the original framework of Cunha and Heckman (2008) to model the joint evolution of health, cognitive skills, and socio-emotional skills. Therefore, we are able to investigate the role of birth order and the underlying mechanisms on multiple dimensions of child human capital in a single, dynamic model of production technologies. Second, Pavan (2016) is the only study that uses a dynamic factor model to quantify the birth order effects over childhood, but he only models the evolution of child cognitive skills. We extend Pavan's modeling framework to account for the joint evolution of of health, cognitive skills, and socio-emotional skills.

3.3.2 Decomposition of birth order effects

The joint modelling of the production technology (3.2) and the investment choice (3.3) makes it clear that the effects of birth order can be decomposed in different ways. In this section we define and illustrate their computation. With respect to the time span, we distinguish between contemporaneous and cumulative effects. The *contemporaneous effect* of birth order on parental investments, we denote with $\Upsilon_{I,b,s}$, is short-term, arising within period s conditional on the existing stocks of capabilities $\theta_{k,s}$ and covariates (w_s, X_s, HSP_s). Similarly, contemporaneous birth order effects on child capabilities, $\Upsilon_{k,b,s}$, are also short-term and estimated conditional on $(\theta_{\ell,s}, w_s, X_s)$. We also define *cumulative effects* on parental investments, $\Gamma_{I,b,s}$, and child capabilities, $\Gamma_{k,b,s}$, that are accumulated from the initial stage until stage s.

We then decompose each of the two types of effects, contemporaneous and cumulative, into different components: direct, total, and indirect effect.

The *direct effect* is the impact of birth order on $\theta_{k,s+1}$ and I_s which is not mediated by any other inputs and goes through the TFP. The direct effect of birth order on $\theta_{k,s+1}$ is $\gamma_{k,b,s}$ and captures the impact of birth order on the technological efficiency of the production function of $\theta_{k,s+1}$, conditional on existing stocks of capabilities $\theta_{\ell,s}$ and existing parental monetary investments I_s. The direct effect of birth order on I_s is $\gamma_{I,b,s}$ that captures the effects of birth order on investment decision at stage s, conditional on θ_s and HSP_s, in Equation (3.3). The direct effect captures the differences in the efficiency of the production technology and the investment function or changes in the peer environment for children of different birth order. For example, parents may acquire and accumulate knowledge on child rearing over time which in turn makes the capability production technologies and their investment decision for the later-born children more efficient compared to the earlier-born. In addition, those from larger families may face more intense competition or may use more diversifying strategies to compete for the resources (Howe et al., 2002; Sulloway, 1996).

While the direct effects are clearly defined as above, the total effects can be defined in two different ways. The first holds existing stocks of capabilities θ_s fixed and the total effect of birth order on θ_{s+1} and I_s is computed from the production function (3.2) and the investment decision (3.3) within the stage. We call this first total effect *contemporaneous total effect*. The second holds the initial stocks of capabilities θ_1 fixed. In this case, the total effect at each stage s is accumulated from the initial stage to s. We call this second total effect *cumulative total effect*. We provide the expressions and discuss the interpretation of both contemporaneous and cumulative total effects in the subsection below.

Lastly, the *indirect effects* of birth order are mediated by at least one other inputs and correspond to the total effects net of the direct ones. Therefore, we can define two different indirect effects, corresponding to the two types of total effect they are computed from. In both cases, we assign meaning to the total effect by using reduced-form coefficients (Alwin and Hauser, 1975; Fox, 1980; Schmidt, 1982). We illustrate them in details in the subsections below. In the following discussion, we keep covariates (w_s, X_s, HSP_s), the intercepts $\gamma_{k,0,s}$, and $\gamma_{I,0,s}$ implicit without loss of generality.

Contemporaneous total and indirect effects of birth orders

The contemporaneous total effects of birth order at stage s on I_s and $\theta_{k,s+1}$ are shown by re-writing Equation (3.2) and Equation (3.3) in stage-specific reduced forms:

$$\begin{array}{rcl} \Upsilon_{k,b,s} & = & \gamma_{k,b,s} + \gamma_{k,I,s}\gamma_{I,b,s} \\ \Upsilon_{I,b,s} & = & \gamma_{I,b,s} \end{array} . \tag{3.4}$$

In the decomposition (3.5) the cumulative direct effects of birth order are the same as the contemporaneous direct effects in the contemporaneous decomposition (3.4). However, the indirect effects are not the same. The indirect effect arises from the conjunction of investment productivity ($\gamma_{k,I,s} \neq 0$) and the effects of birth order on I_s ($\gamma_{I,b,s} \neq 0$), keeping $\theta_{\ell,s}$ fixed. The term $\gamma_{k,I,s}$ characterizes how much of the effects of birth order on investment during stage s propagates into differences in child capabilities $\theta_{k,s+1}$. Regarding the contemporaneous effects of BO on I_s, the total effect coincides with the direct effect, captured by $\gamma_{I,b,s}$ and the indirect effect is null.

Cumulative total and indirect effects of birth order

The cumulative total effects of birth order on $\theta_{k,s+1}$ and I_s are accumulated over time starting from the initial stage $s = 1$. Denote by $\Gamma_{I,b,s}$ and $\Gamma_{k,b,s}$ the cumulative total effects of birth order on I_s and $\theta_{k,s+1}$ from the first to current stage. $\Gamma_{k,b,0}$ captures the biological origin of birth order on child capabilities at the initial stage $s = 1$. Recursive substitution of the investment approximation and the capability production using the specifications in Equation (3.3) and (3.2) gives the following representation of the cumulative total effects of birth order on I_s and $\theta_{k,s+1}$:

$$\begin{array}{rcl} \Gamma_{I,b,s} & = & \sum_\ell \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{I,b,s} \\ \Gamma_{k,b,s} & = & \sum_\ell \gamma_{k,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{k,I,s}\Gamma_{I,b,s} + \gamma_{k,b,s} \end{array} \tag{3.5}$$

in which $\Gamma_{\ell,b,s-1}, ..., \Gamma_{\ell,b,2}$ are, in turn, functions of $\Gamma_{\ell,b,s-2}, ..., \Gamma_{\ell,b,1}$, respectively. Note that, in our model, $\Gamma_{\ell,b,1} \equiv \Upsilon_{\ell,b,1} = \gamma_{\ell,b,1}$ and $\Gamma_{I,b,1} \equiv \Upsilon_{I,b,1} = \gamma_{I,b,1}$, that is, the cumulative and contemporaneous total effects during the initial stage coincide and are equal to the direct effects. By backward substitution, the cumulative total birth order effect on parental investments, $\Gamma_{I,b,s}$, is eventually computed from the initial direct effect $\gamma_{I,b,1}$, and the effects of capabilities and birth order on parental investments over time, i.e., $(\gamma_{I,\ell,1}, ..., \gamma_{I,\ell,s})$ and $(\gamma_{I,b,1}, ..., \gamma_{I,b,s})$, respectively. Similarly, the cumulative total birth order effect on child capability $\theta_{k,s+1}$, $\Gamma_{k,b,s}$, is computed from $(\Gamma_{\ell,b,1}, \gamma_{k,\ell,1}, ..., \gamma_{k,\ell,s}, \gamma_{k,I,1}, ..., \gamma_{k,I,s}, \gamma_{k,b,1}, ..., \gamma_{k,b,s})$ with $\Gamma_{\ell,b,1}$ being the cumulative total birth order effect on $\theta_{\ell,2}$. All of these are parameters from the joint model of the production technologies (3.2) and the investment approximation (3.3). The decomposition in (3.5) shows clearly the origins of birth order effects on within-family resource allocations, I_s, and on child capabilities, and is made possible by the joint specification of recursive technologies and investment equations.

To better illustrate the procedure, consider the cumulative total birth order effect on cognitive ability at the third stage, $\theta_{C,3}$. Applying the recursive procedure in Equation (3.5) we obtain:

$$\Gamma_{C,b,3} \quad = \quad \sum_\ell \gamma_{C,\ell,s}\Gamma_{\ell,b,2} + \gamma_{C,I,2}\Gamma_{I,b,2} + \gamma_{C,b,2} \ , \ \ell \in \{H, C, N\}$$

in which we further substitute backward as follows:

$$\begin{array}{rcl} \Gamma_{\ell,b,2} & = & \sum_{\ell'} \gamma_{\ell',\ell,1}\Gamma_{\ell,b,1} + \gamma_{C,I,1}\Gamma_{I,b,1} + \gamma_{C,b,1} \\ \Gamma_{I,b,2} & = & \sum_{\ell'} \gamma_{I,\ell',2}\Gamma_{\ell,b,1} + \gamma_{I,b,1} \end{array}$$

and finally, $\Gamma_{\ell,b,1} = \gamma_{\ell,b,1}$ and $\Gamma_{I,b,1} = \gamma_{I,b,1}$.

Equation (3.5) allows us to speculate further on the cumulative indirect effects. Specifically, the cumulative indirect effect of birth order on I_s, $\sum_{\ell} \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1}$, arises from the conjunction of parental responses to child capabilities in investment decisions ($\gamma_{I,\ell,s} \neq 0$) and the cumulative total effects of birth order on past capabilities ($\Gamma_{\ell,b,s-1} \neq 0$). The degree and direction of parental responsiveness can be either reinforcing ($\gamma_{I,\ell,s} > 0$) or compensating ($\gamma_{I,\ell,s} < 0$) and determines how much of the past cumulative total birth order effects are transmitted to existing parental monetary investments. In the case of multidimensional capability vectors, the total gaps in investments may arise as long as parental monetary investments are responsive to at least one component of the vector and the cumulative birth order total gaps of that capability in the past is non-negligible.

The cumulative indirect effects of birth order on $\theta_{k,s+1}$ are generated from various sources. First, the cumulative total effects in the past ($\Gamma_{\ell,b,s-1} \neq 0$) may work in synergy with (i) the self-productivity of child capabilities ($\gamma_{k,\ell,s} \neq 0$), which include self-effects ($\gamma_{k,k,s}$) and cross-effects ($\gamma_{k,\ell,s}$ with $\ell \neq k$) and/or (ii) the investment productivity ($\gamma_{k,I,s} \neq 0$) to generate the existing cumulative total effects. Second, the birth order gaps may solely come from the contemporaneous total gaps ($\Upsilon_{k,b,s} \neq 0$) of which the decompositions are discussed above[3]. Lastly, the cumulative indirect effetcs of BO on $\theta_{k,s+1}$ consist of the indirect effects from one input to another, making the analysis of the contribution of specific effects transmitted by a particular variable or group of variables to the birth order total gaps obscured.

While it is possible to analyze a number of specific effects, we focus on decomposing further the indirect effects generated through parental monetary investments in conjunction with parents' responses to (i) existing stocks of capabilities and (ii) child birth order. Their computations are shown in Table 3.2. It is worthy to emphasize the importance of the decomposition of birth order effects in our study.

3.3.3 Identification and estimation strategy

3.3.3.1 Identification

The identification of the model specified in Equation (3.2) follows a two-step procedure. In the first step we identify the joint distribution of unobserved capabilities ($\theta_{H,s}, \theta_{C,s}, \theta_{N,s}$) over s. The second step involves the identification of the parameters of the production technologies which allows us to compute the birth order effects.

First, although the capability variables ($\theta_{H,s}, \theta_{C,s}, \theta_{N,s}$) are assumed to be unobserved by the econometrician, we take advantages of observed measures of each capability across different stages to identify their joint distribution over time. Denote the health measures $M_{j,H,s}$ noisy measures of the health stock $\theta_{H,s}$; the test scores $M_{j,C,s}$ and $M_{j,N,s}$ are noisy measures of the cognitive skills and the socio-behavioral skills. Following the approach by Cunha and Heckman (2008), Cunha et al. (2010), and Agostinelli and Wiswall (2016a, 2016b), we use these observed measures to form a measurement system which allows us identify nonparametrically the joint density of all unobservables ($\theta_{H,s}, \theta_{C,s}, \theta_{N,s}$) across $s = 1, ..., S$.

To simplify the computation, we assume a linearly separable measurement system as follows:

$$M_{k,j,s} = \alpha_{1,k,j,s} + \alpha_{2,k,j,s} ln\theta_{k,s} + \mu_{k,j,s} \tag{3.6}$$

in which $\alpha_{1,k,j,s}$ is intercept, $\alpha_{2,k,j,s}$ is factor loading, and the term $\mu_{k,j,s}$ is measurement error. In Equation (3.6), $m_{k,j,s}$ is an age-standardized measure of capability k, not a raw score. For each dimension of capability $\theta_{k,s}$, we use the same anchoring measure $M_{k,1,s}$ over time. In a dynamic setting, this is a critical element that allows us to anchor estimated parameters on a common scale over time (Agostinelli and Wiswall, 2016a, 2016b). In addition, several conditions on the observed measures must be satisfied to use Cunha, Heckman,

[3]Notice that the decomposition (3.5) can also be rewritten in the following form:

$$\Gamma_{k,b,s} \quad = \quad \sum_{\ell} \gamma_{k,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{k,I,s}\left(\sum_{\ell} \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1}\right) + \Upsilon_{k,b,s} \ .$$

Table 3.2: Total, direct, and indirect effects of birth order on investments and capabilities

	Effects on parental monetary investments I_s	**Effects on** Child capability θ_k at $s+1$, $\theta_{k,s+1}$
Contemporaneous total effects	$\Upsilon_{I,b,s} = \gamma_{I,b,s}$	$\Upsilon_{k,b,s} = \gamma_{k,b,s} + \gamma_{k,I,s}\gamma_{I,b,s}$
Direct effects	$\gamma_{I,b,s}$	$\gamma_{k,b,s}$
Indirect effects	n/a	$\gamma_{k,I,s}\gamma_{I,b,s}$
Cumulative total effects	$\Gamma_{I,b,s} = \sum_\ell \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{I,b,s}$	$\Gamma_{k,b,s} = \sum_\ell \gamma_{k,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{k,I,s}\Gamma_{I,b,s} + \gamma_{k,b,s}$
Direct effects	$\gamma_{I,b,s}$	$\gamma_{k,b,s}$
Indirect effects	$\sum_\ell \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1}$	$\sum_\ell \gamma_{k,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{k,I,s}\Gamma_{I,b,s}$
Specific indirect effects through		
... productivity of capabilities	$\sum_\ell \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1}$	$\sum_\ell \gamma_{k,\ell,s}\Gamma_{\ell,b,s-1}$
... productivity of investment	n/a	$\gamma_{k,I,s}\left(\sum_\ell \gamma_{I,\ell,s}\Gamma_{\ell,b,s-1} + \gamma_{k,I,s}\gamma_{I,b,s}\right)$

By backward substitution,$\Gamma_{\ell,b,s-1}, ..., \Gamma_{\ell,b,2}$ are functions of $\Gamma_{\ell,b,s-2}, ..., \Gamma_{\ell,b,1}$, respectively.

and Schennach's (2010) approach. First, if the model contains L unobservables, then one must observe $2L+1$ measures of these variables, a requirement satisfied by the Young Lives data. Second, we must impose the following normalizations on the system (3.6) as follows. (a) The factor mean (in logs) is set to 0, $E(ln\theta_{k,s}) = 0$, for all k and s. (b) The factor loading on one of the measures (say the first measure) of each θ_k is set to 1, i.e., $\alpha_{2,k,1,s} = 1$, for all k and s. (c) There exists a common anchoring measure $m_{k,1,s}$ for each k and in all s. These normalizations, jointly with some regularity conditions, are sufficient to satisfy the nonparametric identification of the joint distributions $\{\theta_{H,s}, \theta_{C,s}, \theta_{N,s}\}_{s=1,..,S}$ in Cunha, Heckman, and Schennach (2010).

In the second step, we establish the identification of parameters in the production technologies. As discussed above, we take advantage of the plausibly exogenous variation in investment incentives generated by the geographic- and time-variant welfare rules under the 1999-2020 National Target Programs in Vietnam to obtain family-specific, stage-specific exclusion restrictions for endogenous parental monetary investment. The welfare rules allow us to take advantage of national housing support programmes (HSP) during the 2000s and 2010s that entitle *every* poor households, living in rural areas to unconditional cash transfers and/or low-interest loan. Moreover, the HSP instruments in this context induce variation in the parental monetary investments across locations and and over time for a single cohort of children, allowing for a tight design to handle the endogeneity of investment.

3.3.3.2 Estimation strategy

Each child is followed from birth to age 12. The interviews are held in four occasions implying a maximum of four observations per child. Information about parental monetary investments is available in three waves. Hence, we assume $s = 1, 2, 3, 4$ with $s = 1$ being the initial stage at birth. We allow for stage-specific parameters in both the production technologies and the investment equations. Within each stage, the model parameters are assumed to be constant. We assume that missing capability measures are missing at random (MAR). However, as explained in the data description, we select the sample such that parental monetary investments and observed covariates X are not missing. We assume that all random variables are drawn from normal distributions, a non-restrictive assumption as noted in the Appendix A11.2. of Cunha, Heckman, and Schennach (2010). The model is then estimated by using the maximum likelihood estimator. We estimate the model using the STATA command SEM, version 14.2 (see StataCorp, 2015).

3.4 Empirical results

3.4.1 Estimation of the production technologies and the investment decisions

Given the large number of estimated parameters, we report the most relevant parameters of Equation (3.2) and Equation (3.3) in Table 3.3. Comprehensive tables with full estimation results are provided in Appendix (Table 7, Table 8, and Table 9). The parameters presented in this section are informative about the role of birth order in the recursive production technologies and essential to compute various birth order effects. It is also important to recall that the coefficient $\gamma_{I,b,s}$ in Equation 3.3 corresponds to three effects: (i), (ii) the contemporaneous total and direct effects, (iii) the cumulative direct effect on parental investments I_s. The coefficient $\gamma_{k,b,s}$ in Equation 3.2 corresponds to the contemporaneous and cumulative direct birth-order effect on child capabilities $\theta_{k,s+1}$.

We start by discussing the estimated parameters of the investment equation in Panel A of Table 3.3. Overall, second-born children receive a lower level of monetary investments than first borns but only from age 2 to age 8. The birth-order gap on investments is strongest from age 2 to 5 with $\gamma_{I,b,1} = -0.212$, suggesting that parents invest about 21.2% less on the second-borns than on the first-born, conditional on the existing stocks of child capabilities and family backgrounds. However, the gaps lower to about 8.5% in the third stage and almost disappear in the fourth stage. Throughout the childhood, parental monetary investments are positively responsive to child existing capabilities, which implies a reinforcing pattern of parental monetary investments. Children with higher stocks of health, cognitive skills and socio-behavioural skills receive significantly higher investments over three stages of childhood.

Table 3.3: Selected parameters of the investment decision and the production technologies

Selected parameters of the investment decision (Eq. 3.3) and the production technologies (Eq. 3.2)

	S=2, age 1-5 coeff. (conf. int.)	**S=3, age 6-8** coeff. (conf. int.)	**S=4, age 9-12** coeff. (conf. int.)
Panel 3.3A. parental monetary investments			
Birth order	-0.212***	-0.085**	0.006
($\Upsilon_{I,b,s} = \gamma_{I,b,s}$,contemporaneous total/direct effect)	(-0.327 -0.097)	(-0.169 -0.002)	(-0.081; 0.092)
Health	0.171***	0.049**	0.0161***
	(0.099; 0.243)	(0.001; 0.097)	(0.019; 0.104)
Cognitive skills		0.195***	0.294***
		(0.133; 0.258)	(0.200; 0.388)
Socio-behavioral skills			0.122***
			(0.028; 0.217)
Panel 3.3B. Health			
Birth order	-0.009	-0.016	-0.108**
($\gamma_{H,b,s}$, contemporaneous/cumulative direct effect)	(-0.091; 0.074)	(-0.113; 0.145)	(-0.197 -0.020)
parental monetary investments	0.053**	0.161***	0.183***
	(0.000; 0.105)	(0.050; 0.272)	(0.102; 0.263)
Health	0.739***	0.181***	0.249***
	(0.688; 0.791)	(0.107; 0.256)	(0.205; 0.294)
Cognitive skills		0.214***	0.327***
		(0.116; 0.313)	(0.227; 0.426)
Socio-behavioral skills			0.094*
			(-0.003; 0.191)
Panel 3.3C. Cognitive skills			
Birth order	-0.032	-0.031	0.006
($\gamma_{C,b,s}$, contemporaneous/cumulative direct effect)	(-0.116; 0.052)	(-0.071; 0.010)	(-0.011; 0.023)
Parental monetary investments	0.188***	0.084***	0.075***
	(0.135; 0.241)	(0.024; 0.144)	(0.025; 0.124)
Health	0.065**	0.059***	-0.007
	(0.012; 0.117)	(0.035; 0.082)	(-0.015; 0.002)
Cognitive skills		0.493***	0.482***
		(0.460; 0.525)	(0.458; 0.505)
Socio-behavioral skills			0.069***
			(0.050; 0.088)
Panel 3.3D. Socio-behavioral skills			
Birth order		-0.036	-0.029
($\gamma_{N,b,s}$, contemporaneous/cumulative direct effect)		(-0.094; 0.022)	(-0.107; 0.049)
parental monetary investments		0.078**	0.059
		(0.010; 0.146)	(-0.015; 0.134)
Health		0.017	-0.01
		(-0.016; 0.051)	(-0.049; 0.030)
Cognitive skills		0.074***	0.136***
		(0.029; 0.120)	(0.048; 0.225)
Socio-behavioral skills			0.144***
			(0.058; 0.230)
N	961	961	961

Table A8 in Appendix A shows full estimation results for the investment equation. Throughout childhood, parental wealth stands out as the most important determinant of investments, surpassing the total effect sizes of all other factors (child capabilities, parental education, HSP eligibility, residence). However, even in the presence of family wealth, parental education and residence fixed effects, welfare eligibility still play a crucial role on parental monetary investment decisions. The test for joint significance of welfare eligibility indicates that we cannot reject that they jointly have effects on parental monetary investments at the 1% level. The bottom panel of Table A8 shows the correlation matrix between the unobserved part of investment and the unobserved part of child capabilities. The correlations are negative and statistically significant, implying that failure to account properly for investment endogeneity would result in downward biased estimates of investment productivity in the production technologies. Furthermore, this would also induce a bias in the estimated contemporaneous and cumulative birth order gaps on child capabilities and investments as we have already discussed in Section 3.3.2.

We now turn to the child capabilities equations, whose estimated parameters are shown in Panel B-D of Table 3.3. The estimates of birth-order dummy for the production function show that second-born siblings are likely to have a less efficient production technology than the first-born. This occurs from age 1 to age 12 for health and socio-emotional skills and from age 1 to age 8 for cognitive ability. The sign reverses only in the last stage of cognition development. However, the direct effects of birth order on the production technology are rather small and statistically insignificant in most stages. The only exception is the fourth stage of health development, during which second-born children have a significantly less efficient production technology, $\gamma_{H,b,3} = -0.108$. The impact of investments on different capability dimensions are always large and significant, but are much larger for health and cognitive ability. While the investment productivity on health is increasing as children grow up, it diminishes across developmental stages for cognitive development.

Panel 3.3B-3.3D contain strong evidence of self-productivity of child capabilities: both self-effects and cross-effects among different elements are significant and sizeable. The self-effects of child health decline with age while cognitive skills are likely to be equally effective at different stages in promoting themselves. Across all stages, the cross-effectiveness of child capabilities is sizeable and statistically significant, highlighting the importance of joint estimating their development process.

Finally, we briefly discuss the findings of the measurement system in Equation (3.6) (Table 7 in Appendix shows the full estimated factor loadings of capability measures). Table 3.5 shows the signal-to-noise ratios of capability measures, which indicates the informational content of capability measure $M_{k,j,s}$ and is calculated as below (Cunha and Heckman, 2008):

$$\iota_j^{ln\theta_{k,s}} = \frac{\alpha_{2,j,k,s}^2 \text{var}(ln\theta_{k,s})}{\alpha_{2,j,k,s}^2 \text{var}(ln\theta_{k,s}) + \text{var}(\mu_{j,k,s})}. \tag{3.7}$$

The closer to unity the ratio the higher informational content of the observed capability measure whereas a ratio closer to zero indicates that the measure is not informative. None of the observed measures provide perfect information about the unobserved capabilities. The PPVT scores in the second stage has the highest signal-to-noise ratio ($\iota_1^{ln\theta_{C,2}} = 0.772$) but about 23 percent of the measure are still contaminated by measurement errors. This result highlights the importance of accounting for measurement errors in child capabilities. Across different dimensions the height-for-age z-scores are most informative about child's health in every stage, while the measures with highest informational content of cognitive skills and socio-behavioral skills changes over time. Measures of cognitive skills are more comparable in their signal-to-noise ratios with all of the measures containing at least more than 20 percent signal. Measures of health and socio-behavioral skills are more diverse. The widest informational gaps of health measures are 59 percentage points in the initial stage between height-for-age z-score and BMI-for-age z-score. The figure for socio-behavioural skills is 51.8 percentage points in the fourth stage between child's schooling aspiration and self-efficacy.

3.4.2 The cumulative birth-order effects and their components

To examine the birth-order gaps and to quantify the contribution of different sources we follow the decomposition analysis discussed in subsection 3.3.2, based on joint estimation of the investment equation (3.3) and the

Table 3.5: Signal-to-noise ratios of capability measures

Variable name	**s=1, age 0-1**	**s=2, age 2-5**	**s=3, age 6-8**	**s=4, age 9-12**
Health				
height-for-age z-score	0.654	0.629	0.513	0.589
bmi-for-age z-score	0.064	0.210	0.256	0.276
carergiver's assessment	0.127	0.061	0.019	0.057
Birth weight	0.093			
Cognitive skills				
PPVT score		0.772	0.517	0.208
CDA score		0.313		
Math score			0.477	0.504
EGRA score			0.396	
Reading score				0.397
Sociobehavioral skills				
Self-efficacy			0.357	0.567
Self-esteem			0.412	0.381
Positivity			0.367	
Aspiration				0.049
Child-parental closeness				0.290
N.	961	961	961	961

production technologies (3.2). Table 3.6 presents three types of cumulative effects of birth order on parental monetary investments and child capabilities with the firstborn as the reference category: the total effects, the direct and indirect effects. For child capabilities, we further decompose the indirect gaps into specific effects materializing through child capability and through parental monetary investments respectively.

Three results emerge clearly from this table[4]. First, second-born children's capabilities are significantly lower than first-born's capabilities throughout childhood, with gaps being most pronounced on cognitive development. Second, until age 5 parental monetary investments account for at least half of the total birth order gaps on child health and cognitive skills. However, when children grow up, the birth order gaps are accumulated mostly through the self-productivity of child capabilities. Third, the birth order effects on parental monetary investments arise not only from the direct parents' responses to child birth order but also indirectly through parents' responses to child existing capabilities.

3.4.2.1 Cumulative birth-order effects on parental monetary investments

Panel 3.6A of Table 3.6 presents the cumulative gaps in parental monetary investments from age 2 to 12. The total gap is most pronounced at the second stage (age 2-5), in which parental monetary investments on the second-born children are on average 21.2 percent lower than those addressed to first-born. As children grow up the total birth order gaps halve to only 10.1 percent in the third stage and become insignificant in the fourth stage, with the effect size being equal to only one-tenth of the initial total gap.

Panel 3.6A also contains the decomposition of the total effects into direct and indirect effects, which further clarifies the mechanisms that drive the observed investment gaps. Note that the direct effects come straightforward from the estimated birth order coefficients, $\gamma_{I,b,s}$, in Equation (3.3) and have been discussed previously in Section 3.4.1. For children having similar stocks of existing capabilities and family background, parents systematically invest less on second borns until age 8. These direct parental responses are independent of child skills and represent the main source of the observed total gaps in parental monetary investments from age 2 to age 8. While the role of direct parents' responses to birth order decreases over time, the last row of Panel 3.6A shows the increasing importance of the indirect effects, which almost double in size from stage 3 to stage 4. That is, birth order affects child development and parents treat children differently based on their capabilities.

Lastly, the results shown in the last column of Panel 3.6A demonstrate the value of our decomposition analysis. The cumulative total gap in investments by birth order is no longer statistically significant and indeed, its magnitude is only one-tenth of the total gap in the second stage. Looking only at these total gaps or the direct effects resulting from Equation (3.3) may lead the analyst to incorrectly conclude that there is no birth oder effect on parental monetary investments. However, our results clearly show that at age 9-12, despite parents' investments no longer respond directly to child birth order, there still exist significant indirect birth order effects which materialize through parents' responses to child capabilities.

While researchers have always been interested in parental monetary and non-monetary investments as a driving force of birth-order gaps on children capabilities, a few empirical studies investigate differences in parents' investing behavior by birth order, such as Price (2008), Monfardini and See (2012), Hotz and Pantano (2015), Pavan (2016), Lehmann et al. (2018), and Black et al. (2018). However, these studies do not analyse the origins of the estimated gaps in parental investment nor they are able to quantify the contribution into the total gaps on multidimensional child capabilities through investment channels as we do in this paper. Among these studies, Pavan (2016) is the closest antecedent to our research. He uses data from the Children of the National Longitudinal Survey of Youth 1979 (CNSLY79) to estimate a dynamic factor models of cognitive development jointly with an appoximation of parental monetary investments. While he finds that the differences in parental monetary investments by birth order account for more than one-half of the birth-order gaps in cognitive skills, he does not attempt to further decompose the investment gaps as we do in this paper. Most recently, Lehmann et al. (2018) also documents systematic differences in parental behavior across birth order. The authors, however, neither have data on child-specific investments across different

[4] We checked that all these results are robust when estimated with samples including one-child, two-child and three-child families and with different specifications of the initial stage.

stages of development, nor do they jointly model the child development process with parental monetary investment decisions as we do.

3.4.2.2 Cumulative birth-order effects on child health

Panel 3.6B presents the cumulative birth-order effects on child health over time. The total gaps in the first row indicate that second-born children generally have lower stocks of health throughout the childhood but the effect varies substantially across stages. Second-born children start their childhood with a relatively small and not significant deficit of approximately 0.2 percent in stocks of child health (or 0.02 standard deviations on height-for-age z-scores) which remains almost unchanged and not significant until age 8. These gaps however widen dramatically in the 9-12 age range, to about 15 percent of the health factor (or 0.15 standard deviations of height-for-age z-scores).

In the next rows we show the decomposition of the total birth order effects: (i) the direct effects of birth order on stage-specific production efficiency, (ii) the indirect effects of birth order which we further break down into two components: the indirect effects through self-productivity of child capability and through parental monetary investments. In sum, the results suggest that until age 8, the birth order gaps in child health are mainly driven by the differences in parental monetary investments. However, the effect occurring through parental monetary investment become negligible and from age 9 to age 12, confirming the previous result that parental monetary investments are not responsive to birth order per se (column 4 Panel 3.6A). Indeed, at this stage, the health disadvantage of second-borns are driven by their less efficient health production technology and by the birth order gaps accumulated in existing stocks of child capabilities, which are transmitted to child health in the next stage through self-productivity.

3.4.2.3 Cumulative birth-order effects on cognitive skills

In contrast to the increasing disadvantage in health of the second-borns, our model implies significant but decreasing cumulative total effects on child cognitive skills by birth order across developmental stages. Panel 3.6C of Table 3.6 shows that second-born children persistently have lower stocks of cognitive skills than the first-borns. The gap is about 7.6 percent of the cognitive factor from age 2 to age 8 (or equivalently, approximately 0.076 and 0.078 standard deviations of the standardized PPVT scores and PIAT-Math scores, respectively). The total gap declines by one-half to 3.5 percent of the stocks of cognitive skills in the fourth stage (0.035 and 0.06 standard deviations of the standardized PPVT scores and PIAT-Math scores, respectively). Comparing to the two closely related studies in the U.S context by Pavan (2016) and Lehmann et al. (2018), we find a similar pattern of declining effects on cognitive skills. However, the birth order effects on Vietnamese children appear to be substantially lower[5].

The lower rows of Panel 3.6C show the contribution of different sources to the total birth order gap in cognitive skills. Similar to the pattern spotted for health, we find that the importance of the investment channel declines substantially over time, mimicking closely the decreasing detrimental effect of birth order on parental monetary investments. However, differently from the findings on health, the accumulation of birth order effects through self-productivity plays here a crucial role, representing from 50 to 100 percent of the cumulative total gap.

3.4.2.4 Cumulative birth-order effects on socio-behavioral skills

Panel 3.6D shows the cumulative birth order effects on socio-behavioral skills. The total effects are fairly large compared to those on health and cognitive skills. We find that the stock of socio-behavioral skills of the second-borns is about 5 percent lower than that of first-borns, even if they start with similar level of stock of initial capabilities (or 0.05 standard deviations of the standardized self-efficacy scores). These effects remain almost quantitatively unchanged from the third to the fourth stage, although we cannot reject the null

[5] Both studies, using different econometric techniques, document quantitatively large gaps in cognitive skills between the first and second borns, ranging from 0.1 to 0.2 standard deviations of test scores (PIAT-Math scores).

hypothesis that the gaps are equal to zero. As the decomposition results show in the lower rows, this is likely driven by the lack of statistical power of the direct birth order effects on socio-behavioral skills technology efficiency.

Our results for socio-behavioural skills are quantitatively similar to the birth order effects found in Black et al. (2018) and Lehmann et al (2018). However, these two studies control for child cognitive skills, despite the latter are an outcome variable on which birth order also has strong negative effects. By contrast, the dynamic modelling approach we adopt in this paper allows us to take into account the birth order effects arising through capability self-productivity, and to quantify the contribution of this channel.

3.4.2.5 Visual inspection of the cumulative versus contemporaneous birth order effects

Figure 3.1 plots the cumulative and contemporaneous total effects of birth order. These figures clearly demonstrates the value of our decomposition analysis that depending on the questions of interests, i.e., whether we are interested in the long-term or the short-term effects of birth order on different outcomes, the answers are in general remarkably different. Indeed, this distinction has not been clarified in the previous studies.

The curves of cumulative total effects always lie below those of contemporaneous effects at any stages of development, regardless of the outcomes of interest, and the differences between them gradually enlarge. Indeed, while the two types of effects in general follow a similar direction (increasing or decreasing), they can also be remarkably contradictory as in the case of investments and cognitive skills at stage 4. For example, looking only at the contemporaneous total effect on cognitive skills from age 9-12 (the dashed line in Figure 3.1c), which is estimated by conditioning on the existing stock of child capabilities, may lead one to conclude that birth order is unlikely to have any effects. However, the cumulative total effect (the solid line in Figure 3.1c) clearly indicates cognitive disadvantages of second-born children.

In addition, the four sub-figures from 3.1a to 3.1d follow two clearly distinct patterns of birth order effects over time. Specifically, the birth order gaps in parental investments and cognitive skills narrow down over time. This reflects the results discussed in subsection 3.4.2 that parental investments are most responsive to children's cognitive skills and that parental investments are highly productive in the production function of child cognition. In contrary, second-born children are lagged behind more and more in terms of their health, reflecting the previous finding that the production function of health for second-born children becomes significantly less efficient in stage 4.

3.5 Discussion and concluding remarks

In this paper, we focus on two children families in Vietnam and show that the negative relation between birth order and child health, cognitive skills and socio-behavioral skills starts very early in life. Later-born children are in general less healthy than the first-borns with the gaps enlarging significantly as children grow up. By age 12, the stock of health of second-born children is about 12 percent lower than that of the first borns. By contrast, the birth order gaps in cognitive skills, are sizeable and significant from the first five years of life until adolescence, but they gradually decline over time. We find evidence that socio-behavioral abilities are lower for later-born children from age 6 to age 12.

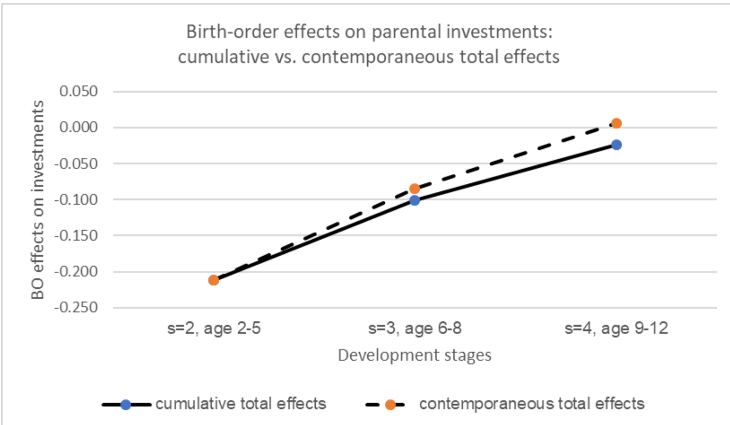
We go beyond the previous studies on birth order effects and propose a decomposition which identifies and quantifies the contribution of different mechanisms to the total birth order gaps in child capabilities. Crucial to this achievement is the specification of a joint model of dynamic production technologies and parental monetary investment decision, similar to the model used by Cunha and Heckman (2008) and Cunha et al. (2010). The results from the decomposition analysis show that there exists three possible channels that work in synergy to generate the birth order capability gaps: (i) the birth order effect on the production technological efficiency, (ii) the birth order effect on parental monetary investments, (iii) the self-productiveness of child capabilities. While the first mechanism is found to be relevant only in health production technology in the last developmental stage we consider (age 9-12), the contributions of the second and third mechanism are estimated to be sizeable at different stages of development.

Table 3.6: Cumulative birth-order effects on investments and capabilities

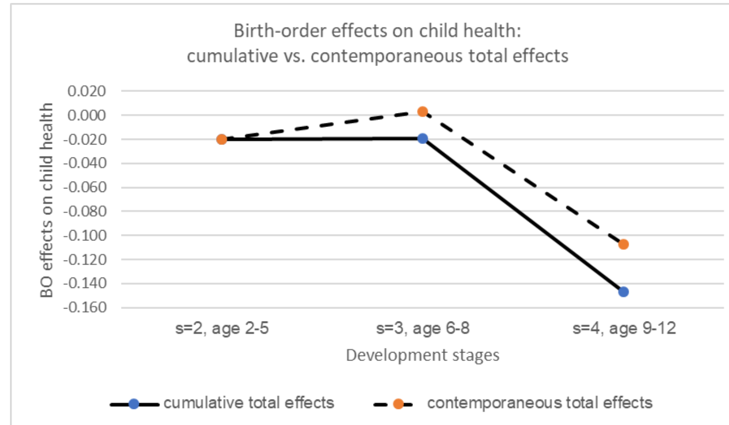
Cumulative effects / Development stages	**s=2, age 2-5** coeff. (conf. int.)	**s=3, age 6-8** coeff. (conf. int.)	**s=4, age 9-12** coeff. (conf. int.)
Panel 3.6A. parental monetary investments			
Cumulative total effects	-0.212***	-0.101**	-0.024
	(-0.327; -0.097)	(-0.185; -0.015)	(-0.113; 0.065)
Direct effects	-0.212***	-0.085**	0.006
	(-0.327; -0.097)	(-0.169; -0.002)	(-0.081; 0.092)
Cumulative indirect effects	n/a	-0.015*	-0.029***
(through parental responses to capabilities)		(-0.033; 0.003)	(-0.053; -0.006)
Panel 3.6B. Health			
Cumulative total effects	-0.020	-0.019	-0.147***
	(-0.102; 0.062)	(-0.150; 0.112)	(-0.246; -0.048)
Direct effects	-0.009	0.016	-0.108***
	(-0.091; 0.074)	(-0.113; 0.145)	(-0.197; -0.020)
Cumulative indirect effects	-0.011*	-0.035**	-0.038
	(-0.024; 0.001)	(-0.067; -0.004)	(-0.086; 0.009)
Specific cumulative indirect effects			
... through capability self-productivity	n/a	-0.019	-0.034*
		(-0.044; 0.006)	(-0.076; 0.007)
... through investment productivity	-0.011*	-0.016*	-0.004
	(-0.024; 0.001)	(-0.034; 0.001)	(-0.021; 0.012)
Panel 3.6C. Cognitive skills			
Cumulative total effects	-0.072*	-0.076***	-0.035**
	(-0.157; 0.013)	(-0.135; -0.017)	(-0.070; -0.001)
Direct effects	-0.032	-0.031	0.006
	(-0.116; 0.052)	(-0.071; 0.010)	(-0.011; 0.023)
Cumulative indirect effects	-0.040***	-0.045**	-0.042***
	(-0.064; -0.015)	(-0.089; -0.001)	(-0.073; -0.011)
Specific cumulative indirect effects			
... through capability self-productivity	n/a	-0.037*	-0.040***
		(-0.079; 0.006)	(-0.069; -0.011)
... through investment productivity	-0.040***	-0.008*	-0.002
	(-0.064; -0.015)	(-0.018; 0.001)	(-0.009; 0.005)
Panel 3.6D. Socio-behavioral skills			
Cumulative total effects		-0.05*	-0.048
		(-0.107; 0.008)	(-0.127; 0.031)
Direct effects		-0.036	-0.029
		(-0.094; 0.022)	(-0.107; 0.049)
Cumulative indirect effects		-0.014**	-0.019 **
		(-0.026; -0.001)	(-0.035; -0.003)
Specific cumulative indirect effects			
... through capability self-productivity		-0.006	-0.017***
		(-0.013; 0.002)	(-0.031; -0.003)
... through investment productivity		-0.008*	-0.001
		(-0.017; 0.002)	(-0.007; 0.004)
N.	961	961	961

Figure 3.1: Cumulative and contemporaneous total effects of birth order

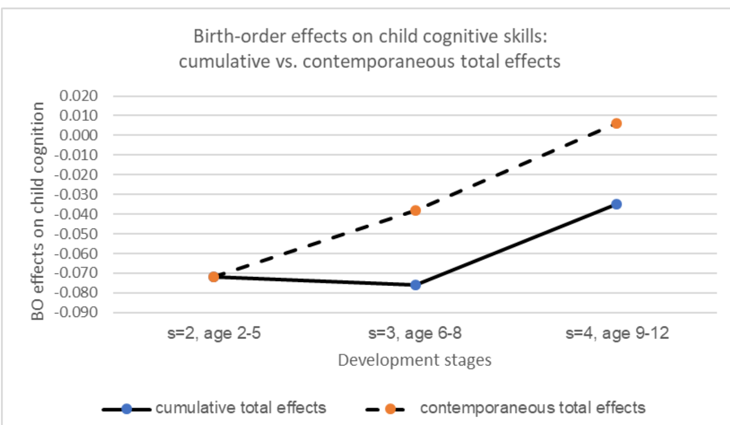
(a) Cumulative and contemporaneous total effects of birth order on parental monetary investments



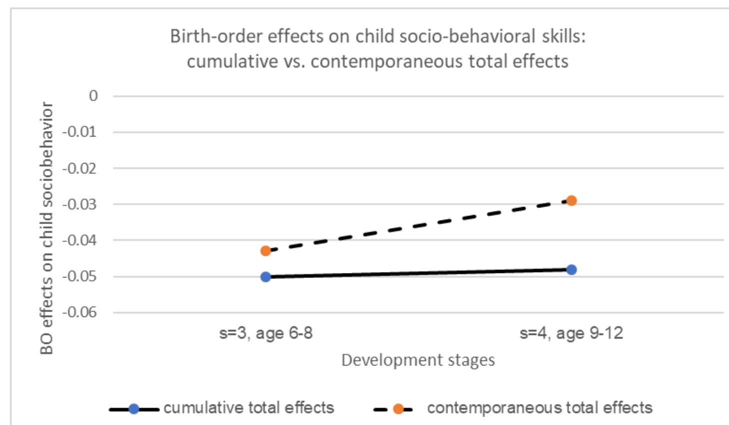
(b) Cumulative and contemporaneous total effects of birth order on child health



(c) Cumulative and contemporaneous total effects of birth order on child cognitive skills



(d) Cumulative and contemporaneous total effects of birth order on child socio-emotional skills



Among the three mechanisms, the birth order effects on parental monetary investments appear to be the most sustained and driving force of the capability gaps in early life. We find that the difference in parental monetary investments by birth order accounts for about one-half of the gaps in health from age 2 to age 8 and in cognitive skills in the first five years of life. However, the contribution to the capability gap occurring through parental monetary investments decreases considerably and become insignificant as children reach age 9-12. This reflects our finding that parents make investment decisions in direct response to child birth order only until age 8. After age 8, the total gap in parental monetary investments by birth order becomes negligible and statistically insignificant. This result on the contribution of the investment channel is largely in line with the findings in the U.S context by Pavan (2016) and Lehmann et al. (2018).

Taken together, our results suggest that the negative relation between birth order and child capabilities in Vietnam can be explained mainly by: (i) differential parental monetary investments by birth order, especially in early life, and (ii) the accumulation of birth order effects through the self-productivity of child capabilities.

Bibliography

[1] Adler, Alfred. "Characteristics of the first, second, and third child." Children 3.5 (1928): 14.

[2] Anderson, M., 2008. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” Journal of the American Statistical Association 103(484):1481–95.

[3] Argys, L. M., D. I. Rees, S. L. Averett, and B. Witoonchart (2006).”Birth order and risky adolescent behavior”. Economic Inquiry 44 (2) (APR): 215-33.

[4] Baker, Paula C., Canada K. Keck, Frank L. Mott, and Stephen V. Quinlan. 1993. The NLSY Child Handbook 1993: A Guide and Resource Document for the NLS of Youth—Child Data. Columbus, OH: The Ohio State University, Center for Human Resource Research.

[5] Behrman, Jere R., and Paul Taubman. 1986. “Birth Order, Schooling, and Earnings.” Journal of Labor Economics 4(3):121–45.

[6] Belfield, Clive R., and Inas Rashad Kelly. 2010. “The Benefits of Breastfeeding Across the Early Years of Childhood.” NBER Working Paper 16496, National Bureau of Economic Research, Inc.

[7] Birdsall, Nancy. 1991. “Birth Order Effects and Time Allocation.” Research in Population Economics 7:191–213.

[8] Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education.” Quarterly Journal of Economics 120(2):669–700.

[9] ———. 2007. “Older and Wiser? Birth Order and IQ of Young Men.” NBER Working Paper 13237, National Bureau of Economic Research.

[10] ———. 2015. “Healthy(?), Wealthy, and Wise: Birth Order and Adult Health.” NBER Working Paper 21337, National Bureau of Economic Research.

[11] Black, S.E., Grönqvist, E. and Öckert, B., 2018. Born to lead? The effect of birth order on noncognitive abilities. Review of Economics and Statistics, 100(2), pp.274-286.

[12] Bleske-Rechek, April, and Jenna A. Kelley. 2014. “Birth Order and Personality: AWithin-Family Test Using Independent Self-Reports from Both Firstborn and Later born Siblings.” Personality and Individual Differences 56:15–18.

[13] Booth, Alison L., and Hiau Joo Kee. 2009. “Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment.” Journal of Population Economics 22(2):367–97.

[14] Brenøe, Anne A., and Ramona Molitor. 2015. “Birth Order and Health of Newborns: What Can We Learn from Danish Registry Data?” CINCH Working Paper 13.

[15] Buckles, Kasey, and Shawna Kolka. 2014. “Prenatal Investments, Breastfeeding, and Birth Order.” Social Science & Medicine 118:66–70.

[16] Buckles, Kasey S., and Elizabeth L. Munnich. 2012. “Birth Spacing and Sibling Outcomes.” Journal of Human Resources 47(3):613–42.

[17] Cunha, Flavio, and James J. Heckman. 2007. “The Economics of Human Development: The Technology of Skill Formation.” The American Economic Review 97(2):31–47.

[18] ———. 2008. “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation.” Journal of Human Resources 43(4):738–82.

[19] Cunha, F., Heckman, J.J. and Schennach, S.M., 2010. Estimating the technology of cognitive and noncognitive skill formation. Econometrica, 78(3), pp.883-931.

[20] Currie, Janet, and Jeffrey Grogger. 2002. “Medicaid Expansions and Welfare Contractions: Offsetting Effects on Prenatal Care and Infant Health?” Journal of Health Economics 21(2): 313–35.

[21] De Haan, Monique, Erik Plug, and José Rosero. 2014. “Birth Order and Human Capital Development Evidence from Ecuador.” Journal of Human Resources 49(2):359–92.

[22] Deming, David. 2009. “Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start.” American Economic Journal: Applied Economics 1(3):111–34.

[23] Ejrnæs, M., and Claus C. Pörtner. 2004. “Birth Order and the Intrahousehold Allocation of Time and Education.” Review of Economics and Statistics 86(4):1008–19.

[24] Ernst, C., & Angst, J. (2012). “Birth order: Its influence on personality”. Springer Science & Business Media.

[25] Fingerhut, L.A., J.C. Kleinman, and J.S. Kendrick. 1990. “Smoking Before, During, and After Pregnancy.” American Journal of Public Health 80(5):541–4.

[26] Fryer, Ronald, and Steven Levitt. 2004. “Understanding the Black-White Test Score Gap in the First Two Years of School.” Review of Economics and Statistics 86(2):447–64.

[27] Galton, F. (1875). English men of science: Their nature and nurture. D. Appleton.

[28] Hanushek, Eric A. 1992. “The Trade-off between Child Quantity and Quality.” Journal of Political Economy 100(1):84–117.

[29] Heckman, James J., and Dmitri V. Masterov. 2007. “The Productivity Argument for Investing in Young Children.” Applied Economic Perspectives and Policy 29(3):446–93.

[30] Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” Journal of Labor Economics 24(3):411–82.

[31] Hofferth, Sandra L. 2009. “Changes in American Children’s Time—1997 to 2003.” Electronic International Journal of Time Use Research 6(1):26–47.

[32] Horwood, L. John, and David M. Fergusson. 1998. “Breastfeeding and Later Cognitive and Academic Outcomes.” Pediatrics 101(1):e9.

[33] Hotz, V. Joseph, and Juan Pantano. 2015. “Strategic Parenting, Birth Order and School Performance.” Journal of Population Economics 28(4):911–36.

[34] Juhn, Chinhui, Yona Rubinstein, and C. Andrew Zuppann. 2015. “The Quantity–Quality Tradeoff and the Formation of Cognitive and Non-cognitive Skills.” Working Paper 21824, National Bureau of Economic Research.

[35] Kantarevic, Jasmin, and Stéphane Mechoulan. 2006. “Birth Order, Educational Attainment, and Earnings.” Journal of Human Resources 41(4):755–77.

[36] Kessler, Daniel. 1991. “Birth Order, Family Size, and Achievement: Family Structure and Wage Determination.” Journal of Labor Economics 9(4):413–26.

[37] Kilburn, Tina R., Hanne-Lise Falgreen Eriksen, Mette Underbjerg, Poul Thorsen, Erik Lykke Mortensen, Nils Inge Landrø, Leiv S. Bakketeig, Jakob Grove, Claus Sværke, and Ulrik Schiøler Kesmodel. 2015. "Low to Moderate Average Alcohol Consumption and Binge Drinking in Early Pregnancy: Effects on Choice Reaction Time and Information Processing Time in Five-Year-Old Children." PLoS ONE 10(9):e0138611.

[38] Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." Econometrica 75(1):83–119.

[39] Lang, Kevin, and Carlos E. Sepulveda. 2007. "The Black–White Test Score Differential." Boston University Working Paper.

[40] Lehmann, J.Y.K., Nuevo-Chiquero, A. and Vidal-Fernandez, M., 2018. The early origins of birth order differences in children's outcomes and parental behavior. Journal of Human Resources, 53(1), pp.123-156.

[41] Lewis, Caroline T., T.J. Mathews, and Robert L. Heuser. 1996. "Prenatal Care in the United States, 1980–94." Vital and Health Statistics. Series 21, Data on Natality, Marriage, and Divorce 54. Washington, DC: National Center for Health Statistics

[42] Monfardini, Chiara, and Sarah Grace See. 2016. "Birth Order and Child Cognitive Outcomes: An Exploration of the Parental Time Mechanism." Education Economics 24(5):481–95.

[43] Mott, Frank L. 1991. "Developmental Effects of Infant Care: the Mediating Role of Gender and Health." Journal of Social Issues 47(2):139–58.

[44] Oddy, Wendy H., Garth E. Kendall, Eve Blair, Nicholas H. De Klerk, Fiona J. Stanley, Louis I. Landau, Sven Silburn, and Stephen Zubrick. (2003). "Breastfeeding and Cognitive Development in Childhood: A Prospective Birth Cohort Study." Paediatric and Perinatal Epidemiology 17(1):81–90.

[45] Parish, William L., and Robert J. Willis. (1993). "Daughters, Education, and Family Budgets Taiwan Experiences." Journal of Human Resources 28(4):863–98.

[46] Pavan, Ronni. (2016). "On the Production of Skills and the Birth Order Effect." Journal of Human Resources 51(3):699–726.

[47] Price, Joseph. (2008). "Parent–Child Quality Time." Journal of Human Resources 43(1):240–65.

[48] Todd, Petra E., and Kenneth I. Wolpin. (2007). "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." Journal of Human Capital 1(1):91–136.

[49] Richiardi, L., O. Akre, M. Lambe, F. Granath, Scott M. Montgomery, and Anders Ekbom (2004). "Birth Order, Sibship Size, and Risk for Germ-Cell Testicular Cancer", Epidemiology, Vol. 15(3): 323-329

[50] Rees, D. I., E. Lopez, S. L. Averett, and L. M. Argys (2008). "Birth order and participation in school sports and other extracurricular activities" Economics of Education Review 27(3) (JUN): 354-62.

[51] Schooler, C. (1972). Birth order effects: Not here, not now. Psychological Bulletin, 78(3), 161.

[52] StataCorp. (2015). Stata: Release 14. Statistical Software. College Station, TX: StataCorp LLC.

[53] Sulloway, F. J. (1996). Born to rebel: Birth order, family dynamics, and creative lives. Pantheon Books.

[54] Zajonc, Robert B. (1976). "Family Configuration and Intelligence: Variations in Scholastic Aptitude Scores Parallel Trends in Family Size and the Spacing of Children." Science 192(4236):227–36.

Appendix

A.3.1 Estimated factor loadings of the measurement system

Table 7: Factor loadings of capability measures

	s=1, age 0-1 factor loadings (conf. int.)	**s=2, age 2-5** factor loadings (conf. int.)	**s=3, age 6-8** factor loadings (conf. int.)	**s=4, age 9-12** factor loadings (conf. int.)
Health				
height-for-age z-score	1	1	1	1
	(normalized)	(normalized)	(normalized)	(normalized)
bmi-for-age z-score	0.248***	0.645***	0.919 ***	0.797***
	(0.171; 0.325)	(0.565; 0.726)	(0.817; 1.021)	(0.702; 0.892)
carergiver's assessment	0.248***	0.211 ***	0.126***	0.180***
	(0.171; 0.325)	(0.161; 0.262)	(0.082; 0.17)	(0.134; 0.226)
Birth weight	0.312			
	(0.239; 0.385)			
Cognitive skills				
PPVT score		1	1	1
		(normalized)	(normalized)	(normalized)
CDA score		0.625***		
		(0.519; 0.73)		
PIAT-Math score			1.025***	1.816***
			(0.905; 1.144)	(1.499; 2.132)
EGRA score			0.950 ***	
			(0.835; 1.066)	
Reading score				1.637***
				(1.346; 1.928)
Sociobehavioral skills				
Self-efficacy			1	1
			(normalized)	(normalized)
Self-esteem			1.049***	0.818***
			(0.846; 1.252)	(0.682; 0.955)
Positivity			0.806***	
			(0.653; 0.960)	
Aspiration				0.195***
				(0.122; 0.268)
Child-parental closeness				0.713***
				(0.585; 0.84)
N.	961	961	961	961

A.3.2 Full estimated parameters of the production technologies and parental monetary investment equations

Table 8: Estimated parameters of the investment equation

	S=2, age 1-5 coeff. (conf. int.)	**S=3, age 6-8** coeff. (conf. int.)	**S=4, age 9-12** coeff. (conf. int.)
Health	0.153***	0.046*	0.054**
	(0.082; 0.224)	(-0.002; 0.094)	(0.011; 0.096)
Cognitive skills		0.187***	0.276***
		(0.125; 0.250)	(0.181; 0.370)
Sociobehavioral skills			0.110**
			(0.016; 0.203)
HSP eligibility	0.069*	0.048*	0.081**
	(-0.013; 0.151)	(-0.009; 0.106)	(0.003; 0.159)
Wealth index	2.416***	2.013***	2.270***
	(1.946; 2.885)	(1.574; 2.452)	(1.762; 2.779)
Caregiver's education	0.303***	0.115***	0.244***
	(0.226; 0.381)	(0.059; 0.170)	(0.186; 0.302)
Birth order (second-born =1)	-0.195***	-0.079*	0.014
	(-0.309; -0.081)	(-0.163; 0.004)	(-0.072; 0.100)
Log(TFP)	-3.610***	-2.026***	-3.066***
	(-3.968; -3.252)	(-2.416; -1.636)	(-3.407; -2.726)
Correlation matrix between parental monetary investment and			
Health	-0.055***	-0.049***	-0.067***
	(-0.091; -0.018)	(-0.081; -0.016)	(-0.113; -0.022)
Cognitive skills	-0.054***	-0.155***	-0.357***
	(-0.090; -0.018)	(-0.257; -0.053)	(-0.568; -0.146)
Sociobehavioral skills		-0.108***	-0.076***
		(-0.180; -0.036)	(-0.128; -0.025)

Table 9: Estimated parameters of the health technology

	Health			**Cognitive skills**			**Socio-behavioral skills**	
	S=2, age 1-5 coeff. (conf. int.)	S=3, age 6-8 coeff. (conf. int.)	S=4, age 9-12 coeff. (conf. int.)	S=2, age 1-5 coeff. (conf. int.)	S=3, age 6-8 coeff. (conf. int.)	S=4, age 9-12 coeff. (conf. int.)	S=3, age 6-8 coeff. (conf. int.)	S=4, age 9-12 coeff. (conf. int.)
Health	0.740***	0.177***	0.246***	0.061**	0.058***	-0.007	0.016	-0.008
	(0.688; 0.792)	(0.103; 0.251)	(0.202; 0.291)	(0.008; 0.113)	(0.034; 0.081)	(-0.015; 0.002)	(-0.018; 0.049)	(-0.047; 0.031)
Cognitive skills		0.204***	0.321***		0.490***	0.480***	0.071***	0.138***
		(0.105; 0.303)	(0.222; 0.421)		(0.458; 0.522)	(0.457; 0.503)	(0.026; 0.116)	(0.049; 0.226)
Sociobehavioral skills			0.089*			0.069***		0.146***
			(-0.008; 0.186)			(0.050; 0.088)		(0.060; 0.232)
parental monetary investments	0.054**	0.156***	0.180***	0.182***	0.084***	0.076***	0.078**	0.063*
	(0.002; 0.107)	(0.046; 0.266)	(0.100; 0.260)	(0.129; 0.235)	(0.025; 0.143)	(0.028; 0.125)	(0.011; 0.145)	(-0.011; 0.138)
Wealth index	0.504***	0.394	0.168	0.812***	0.145	-0.140**	0.253*	0.232
	(0.176; 0.832)	(-0.156; 0.943)	(-0.280; 0.616)	(0.480; 1.144)	(-0.047; 0.336)	(-0.264; -0.016)	(-0.007; 0.512)	(-0.167; 0.630)
Caregiver's education	0.031	0.06	0.099***	0.124***	0.063***	-0.004	0.018	-0.050*
	(-0.027; 0.089)	(-0.027; 0.147)	(0.036; 0.162)	(0.066; 0.183)	(0.036; 0.091)	(-0.021; 0.012)	(-0.021; 0.058)	(-0.106; 0.006)
Birth order	-0.006	0.02	-0.105**	-0.026	-0.028	0.006	-0.035	-0.031
(second-born =1)	(-0.089; 0.076)	(-0.109; 0.149)	(-0.194; -0.016)	(-0.110; 0.057)	(-0.068; 0.013)	(-0.011; 0.023)	(-0.093; 0.023)	(-0.109; 0.047)
Log(TFP)	-0.286**	-0.385**	-0.313*	-0.186	-0.165**	0.156**	-0.045	0.145
	(-0.529; -0.042)	(-0.746; -0.024)	(-0.666; 0.041)	(-0.432; 0.060)	(-0.305; -0.024)	(0.009; 0.304)	(-0.225; 0.136)	(-0.174; 0.463)
Residence dummies	Y	Y	Y	Y	Y	Y	Y	Y
N	961	961	961	961	961	961	961	961