



Essays in the Evaluation of Human Capital  
Investment Policies

Italo López García

*Supervisors:* Prof. Orazio Attanasio, Prof. Costas Meghir

Thesis submitted to the Department of Economics for the degree of

Doctor of Philosophy (P.hD)  
University College of London

Date of submission: 13/08/14

## **Abstract**

In this thesis dissertation, I study which factors drive human capital investments at different stages of the life-cycle by using structural dynamic behavioral models, and what we can learn from these models in order to design better labor and education policies in the long-term. Over the three chapters of this work I assess policy relevant questions that are either related to both developing and developed countries, and I show how these methodologies can be used in conjunction with other more traditional approaches to perform two types of policy evaluation: the assessment of existing policies, or ex-post policy evaluation, and the prediction of economic behavior under policies that have not been yet implemented, or ex-ante policy evaluation. My work has two main goals. The first goal is methodological. I show the gains of structural modeling in understanding the mechanisms behind human capital investments, for example the disentanglement of preferences, returns and expectations, and the importance of dynamics. I also show how these models can be complemented and even better identified when they are combined with experimental data. The second goal is to answer some relevant economic questions for which there are still no answers. On one hand, I study the determinants of labor informality and self-employment in developing countries disentangling the role of comparative advantage and labor market segmentation on labor informality. On another hand, I study the determinants of parental investments in children and their effects in child development, first emphasizing the role of parental income and financial constraints, and then focusing on less investigated factors like parental beliefs and attitudes towards child-rearing.

# Funding acknowledgments

This thesis has been made possible through the scholarship “Programa de Formación de Capital Humano Avanzado Becas Chile” from The National Commission for Science and Technology CONICYT, Government of Chile.

# Acknowledgments

I would like to thank my supervisors, Professor Costas Meghir, Professor Orazio Attanasio and Professor Pedro Carneiro, for their guidance and support throughout the work of this thesis. I have learnt so much from all of them, not only about economics. A special thanks to Emanuela Galasso, my co-author and boss in one of the chapters of this thesis. Her support was invaluable.

I want to thank to thank the Department of Economics at UCL, which has offered me the opportunity to grow as a person and researcher. I am deeply deeply thankful to all my PhD friends and colleagues that walked this life with me over the past 6 years. Lucia, Florian, Tibo, Luigi, Jan, Jaime, Pamela, Andrés, and many others, thanks!

I owe special thanks to my friends in London and my family in Chile, for the continuous emotional support, in particular at difficult times.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Human capital and labor informality</b>	<b>6</b>
2.1	Introduction . . . . .	6
2.2	Background and Related Literature . . . . .	10
2.3	Data and institutional framework . . . . .	13
2.4	The Model . . . . .	20
2.4.1	The State Space . . . . .	22
2.4.2	Flow Utilities . . . . .	23
2.4.3	The Wage Offer . . . . .	25
2.4.4	Uncertainty . . . . .	25
2.4.5	Recursive problem . . . . .	26
2.4.6	Mobility . . . . .	27
2.5	Model Solution and Estimation . . . . .	28
2.5.1	Solution Method . . . . .	28
2.5.2	Model Identification . . . . .	29
2.5.3	Estimation . . . . .	30
2.6	Results . . . . .	34
2.6.1	Goodness of Fit . . . . .	35
2.6.2	Structural Estimates . . . . .	39
2.7	Studying Segmentation and Dynamics . . . . .	42
2.7.1	Are labor markets segmented? . . . . .	42
2.7.2	The role of dynamics and labor market expectations . . . . .	44
2.8	Concluding Remarks . . . . .	49
<b>3</b>	<b>Does the Timing of Parental Income Matter?</b>	<b>53</b>
3.1	Introduction . . . . .	53
3.2	Data . . . . .	59
3.3	Methods . . . . .	60

3.3.1	Empirical Strategy . . . . .	60
3.3.2	Multivariate Local Linear Regression . . . . .	62
3.4	Results . . . . .	64
3.4.1	Descriptive Statistics . . . . .	64
3.4.2	Parametric Estimates . . . . .	65
3.4.3	Semi Parametric Estimates . . . . .	69
3.4.3.1	Schooling Attainment . . . . .	70
3.4.3.2	Log Earnings at age 30 . . . . .	73
3.4.3.3	High School Drop Out and College Attendance . . . . .	73
3.4.3.4	Other Outcomes . . . . .	76
3.5	Tackling Endogeneity . . . . .	76
3.5.1	Maternity leave choices . . . . .	77
3.5.2	Heterogeneous age-earning profiles . . . . .	78
3.5.3	Heterogeneous preferences . . . . .	79
3.5.4	Further Robustness checks . . . . .	80
3.6	The Mechanisms . . . . .	81
3.6.1	A simple model . . . . .	82
3.6.2	Solution and Estimation . . . . .	86
3.6.3	Results . . . . .	87
3.6.4	Sensitivity . . . . .	91
3.7	Concluding Remarks . . . . .	92
<b>4</b>	<b>The Role of Beliefs in Parental Investments and Child Development</b>	<b>94</b>
4.1	Introduction . . . . .	94
4.2	The intervention . . . . .	97
4.3	Literature Review . . . . .	100
4.3.1	Beliefs and attitudes about child rearing . . . . .	100
4.4	The Evaluation Design . . . . .	103
4.4.1	Target Population . . . . .	103
4.4.2	Sampling . . . . .	103
4.4.3	Power Calculations . . . . .	104
4.4.4	Measurements . . . . .	106
4.5	Treatment Effects . . . . .	110
4.5.1	Identifying treatment effects . . . . .	111
4.5.2	Decomposing treatment effects . . . . .	112
4.5.3	The latent factor model . . . . .	116

---

4.5.4	A Structural Model of Parental Investments in Children . . . .	121
4.6	Baseline Data . . . . .	127
4.6.1	Target Population . . . . .	127
4.6.2	Data descriptives . . . . .	128
4.6.3	Socioeconomic Gradients . . . . .	131
4.6.4	Stylized Facts . . . . .	134
4.7	Concluding remarks . . . . .	137
<b>5</b>	<b>A Collective Model of Labor Informality and Self-employment</b>	<b>139</b>
5.1	Introduction . . . . .	139
5.2	The Model . . . . .	143
5.2.1	Timing and Decisions . . . . .	145
5.2.2	The Young Life . . . . .	146
5.2.3	The Adult Life . . . . .	147
5.2.4	Recursive Formulation . . . . .	151
5.3	Final Comments . . . . .	154
	<b>References</b>	<b>156</b>
	<b>A Additional figures Chapter 2</b>	<b>166</b>
	<b>B Additional figures Chapter 3</b>	<b>170</b>

# List of Figures

2.1	Mean Informality Rates Males . . . . .	16
2.2	Informality Rates by Education . . . . .	17
2.3	Mean wages by education in two sectors: formal and informal em- ployment . . . . .	18
2.4	Labor market participation by sector . . . . .	19
2.5	Model fit wages Less than High School level . . . . .	35
2.6	Model fit wages High School Degree level . . . . .	36
2.7	Model fit wages College Level . . . . .	36
2.8	Simulated Informality Rates by education . . . . .	37
2.9	Model fit Unemployment/Home Production . . . . .	38
2.10	Model fit Schooling Participation . . . . .	38
3.1	Descriptive Statistics . . . . .	65
3.2	Years of schooling v/s paternal Permanent Income. . . . .	66
3.3	Earnings by age 30 v/s paternal Permanent Income . . . . .	66
3.4	Years of schooling v/s paternal Income age 6-11 . . . . .	67
3.5	Earnings by age 30 v/s paternal Income age 12-17 . . . . .	67
3.6	Earnings by age 30 v/s Paternal Income age 6-11 . . . . .	68
3.7	Earnings by age 30 v/s Paternal Income age 12-17 . . . . .	68
3.8	Schooling Attainment v/s Paternal Income age 6-11. <i>PI</i> and Income age 12-17 fixed at the median . . . . .	71
3.9	Schooling Attainment v/s Paternal Income age 12-17. <i>PI</i> and Income age 6-11 fixed at the median . . . . .	71
3.10	Schooling Attainment v/s Paternal Income age 12-17. <i>PI</i> and Income age 0-5 fixed at the median . . . . .	71
3.11	Earnings by age 30 v/s Paternal Income age 6-11. <i>PI</i> and Income age 12-17 fixed at the median . . . . .	74



3.12 Earnings by age 30 v/s Paternal Income age 12-17. <i>PI</i> and Income age 6-11 fixed at the median . . . . .	74
3.13 Earnings by age 30 v/s Paternal Income age 12-17. <i>PI</i> and Income age 0-5 fixed at the median . . . . .	74
3.14 High School Dropout rates v/s Paternal Income age 6-11. <i>PI</i> and Income age 12-17 fixed at the median . . . . .	75
3.15 College attendance v/s Paternal Income age 6-11. <i>PI</i> and Income age 12-17 fixed at the median . . . . .	75
3.16 Schooling attainment v/s Paternal Income age 6-11 excluding age 0-2, <i>PI</i> and Income age 12-17 fixed at median values . . . . .	77
3.17 Schooling attainment v/s Paternal Income age 12-17 excluding age 0-2, <i>PI</i> and Income age 6-11 fixed at median values . . . . .	77
3.18 Schooling attainment v/s Paternal Income age 12-17 excluding age 0-2, <i>PI</i> and Income age 0-5 fixed at median values . . . . .	78
3.19 Schooling attainment v/s Paternal Income age 6-11 controlling for Fixed Effects, <i>PI</i> and Income age 12-17 fixed at median values . . . .	79
3.20 Schooling attainment v/s Paternal Income age 12-17 controlling for Fixed Effects, <i>PI</i> and Income age 6-11 fixed at median values . . . .	79
3.21 Schooling attainment v/s Paternal Income age 12-17 controlling for Fixed Effects, <i>PI</i> and Income age 0-5 fixed at median values . . . . .	79
3.22 Low birth weight v/s Paternal Income age 6-11, <i>PI</i> and Income age 12-17 fixed at median values . . . . .	81
3.23 Low birth weight v/s Paternal Income age 12-17, <i>PI</i> and Income age 0-5 fixed at median values . . . . .	81
3.24 Simulated schooling attainment and Income age 6-11, <i>PI</i> and Income age 12-17 fixed at median values . . . . .	88
3.25 Simulated schooling attainment and Income age 12-17, <i>PI</i> and Income age 6-11 fixed at median values . . . . .	88
3.26 Simulated schooling attainment and Income age 12-17, <i>PI</i> and Income age 0-5 fixed at median values . . . . .	88
3.27 Simulated Investments and Income age 6-11, <i>PI</i> and Income 12-17 fixed at median values . . . . .	89
3.28 Simulated Investments and Income age 12-17, <i>PI</i> and Income 6-11 fixed at median values . . . . .	89
3.29 Simulated Investments and Income age 12-17, <i>PI</i> and Income 0-5 fixed at median values . . . . .	90

---

3.30	Simulated Savings and Income age 6-11, <i>PI</i> and Income 12-17 fixed at median values . . . . .	90
3.31	Simulated Savings and Income age 6-11, <i>PI</i> and Income 12-17 fixed at median values . . . . .	90
3.32	Simulated Savings and Income age 6-11, <i>PI</i> and Income 12-17 fixed at median values . . . . .	91
4.1	Cost-effectiveness Nadie es Perfecto vs Home Visits . . . . .	99
4.2	Mapping Subjective Probabilities to Conditional expectations through IRT models . . . . .	127
4.3	Income per capita in CLP: i) National, ii) National users of the public system, iii) NEP . . . . .	128
4.4	SES gradients in cognitive skills; a) Language; b) Executive functions	132
4.5	SES gradients in non-cognitive skills . . . . .	133
4.6	SES in parental investments a) Non-cognitive stimulation; b) Cognitive stimulation . . . . .	133
4.7	SES gradients in parental beliefs a) parenting styles; b) perceived self-efficacy . . . . .	134

# Chapter 1

## Introduction

Over the last 20 years, the estimation of structural dynamic programming models have widened the frontiers for empirical research in areas like labor economics and development economics. Heckman and Vytlacil (2005) have argued that structural models have several advantages of addressing more complex questions about individual and household behavior that were prohibitive for reduced-form approaches. First, provided the model is correctly identified, one can estimate policy invariant parameters and perform counterfactual policy evaluation, either by extrapolating reduced-form estimates out of sample, or by evaluating the impacts of policies not yet implemented. Second, we can assess the effect of dynamics by incorporating explicitly modeling the role of expectations in optimal decisions. Third, we can disentangle the different mechanisms driving decisions such as preferences, financial constraints and underlying technologies in a straightforward way, as the model provides a direct link between the economic theory and the data. And finally, in the context of policy evaluation, we can control for unobserved heterogeneity in responses to treatment among observationally identical people, if we account for the fact that using different IV deliver might deliver different treatment parameters.

In this thesis dissertation, I study which factors drive human capital investments at different stages of the life-cycle and which types of labor and education policies enhance human capital accumulation in the long-term, by using structural dynamic behavioral models. Over the three chapters of this work I assess policy relevant questions that are either related to both developing and developed countries, and I show how these methodologies can be used in conjunction with other more traditional approaches to perform two types of policy evaluation: the assessment of existing policies, or ex-post policy evaluation, and the prediction of economic behavior under policies that have not been yet implemented, or ex-ante policy evaluation.

My work has two main goals. The first goal is methodological. By using structural estimation alone, or complemented with more traditional methods, I show the gains of this methodology in understanding the importance of human capital investments over the life-cycle, with a particular focus in developing countries (Todd and Wolpin (2010)). First, by modeling individual and household behavior over the life-cycle, I can exploit the dynamic components of decisions in a transparent way. That is to say, agents have expectations about the future and they are an important part of the current decisions. Second, by fully imposing the Human Capital Investment theory in modeling, I can disentangle the incentives for human capital accumulation driven by preferences, economic returns, and expectations. Third, I show how these tools can be easily complemented with other advanced econometric techniques like non-parametric estimation, gaining a deep understanding of the mechanisms behind the data patterns. And finally, by conducting the experimental evaluation of a large-scale policy, I show how a high degree of complementarity between structural estimation and traditional reduced-form methods for policy evaluation can be exploited, breaking the unnecessary and artificial division in the literature between these two-approaches. In particular, I discuss how experimental data is useful to assess the short-term evaluation of policies, but it's also a powerful source of identification within more sophisticated models simulating the long term effects of those policies.

The second goal is to answer some relevant economic questions about the factors driving human capital investments at different stages of the life-cycle. The first research question is placed on the determinants of human capital investments later in the life-cycle and attempts to address to which extent human capital accumulation and preferences are a determinant of labor informality in Latin America, and the long-term assessments of particular educational and labor policies on the size of the informal sector. The other two research questions are placed in understanding the determinants of parental investments in children. The second research question attempts to understand whether the timing of parental income and financial constraints matter for the human capital accumulation of their children when they become adults in Norway, while the third research question looks more at the effects of non-monetary resources on parental investments in children. By exploiting experimental data from a large scale parenting program in Chile, I assess the role of parental beliefs and attitudes towards child-rearing in parental behavior and early child development.

I divide this thesis dissertation in four chapters, each related to one of the three

research questions. In Chapter 2, I develop a dynamic discrete choice model estimated with rich Chilean longitudinal data, in which individuals jointly decide on their schooling and labor participation, to investigate the extent to which human capital accumulation drives participation in informal labor markets. Labor markets in Chile have been documented as being fairly competitive, and provide a unique setting to test for the comparative advantage approach to informality. This way, I explore three potential mechanisms: First, the importance that individuals assign to wages relative to their valuation of non-wage sector attributes; Second, whether individuals accumulate sector-specific human capital with heterogeneous returns; and Third, the importance of the labor market expectations driving labor and schooling choices. Heterogeneous returns are included by modeling unobserved heterogeneity that determine both schooling and sector-specific productivity.

Some findings are worth being discussed. Model estimates suggest that comparative advantage is a more important source of selection into informal jobs than labor market segmentation. I find that human capital accumulation and preferences for job amenities explain up to 75% of transitions between the informal and the formal sector, and the importance of comparative advantage is increasing in education. Second, I test for the importance of labor market expectations and persistency by simulating the effect of a recently implemented 20% wage subsidy in formal jobs targeted at workers between 19 and 26 years old. I find that the subsidy would not only be effective in decreasing informality among the targeted groups, but the incentives to informality also decrease for younger workers (those between 15 to 18 years old). The reduction in informality rates as a consequence of the subsidy would remain persistent until after the age of 40. Furthermore, I find evidence that human capital differs across sectors. High School returns are larger in the formal sector; there is a wage premium to College in the informal sector; and skills acquired in the formal sector in the form of labor experience are more transferable to the informal sector than the other way around. Preferences are also heterogeneous. More educated individuals assign more importance to non-wage sector attributes, particularly in the formal sector, while less educated individuals value higher wage returns, particularly in the informal sector.

Chapter 3 is based on a co-authored paper with Pedro Carneiro (UCL), Kjell Salvanes (Norwegian School of Economics) and Emma Tominey (York University). In this investigation, we analyze whether the timing of parental income shocks matters for the process of skill accumulation of children beyond the importance of permanent income. Using rich longitudinal administrative data for 500,000 individuals in

Norway, we develop parametric and semi-parametric estimation of how the timing of parental income matters for different outcomes measured when their children are in their early 20s. The main finding of the paper is that, for a given level of permanent parental income, balanced income profiles lead to higher levels of education of the child than income profiles subject to several fluctuations. This is because it is better to smooth investments in children than to suffer large fluctuations (which are a consequence of income shocks). This result also holds for related outcomes like high school dropout and college enrolment. For other outcomes, such as earnings, IQ, or teenage pregnancy, the picture can be slightly different, indicating differences in the production function for different outcomes. In terms of magnitudes, although permanent income is more relevant in the production function of human capital outcomes, the timing of income also matters. Increases in permanent income of £100,000 would increase child earnings by age 30 in 10% and would increase schooling attendance in 0.5 years, while if parents were able to shift the same amount from when the child is in middle childhood (6-11 years old) to early childhood (0-5 years old), earnings by age 30 would increase 5% and schooling attendance would increase by 0.25 years. To interpret our findings we complement the non-parametric approach estimating models of parental investment in children with more than one period of childhood, where we emphasize borrowing constraints, uncertainty about income shocks, and parental preferences about own consumption and the human capital of their children. Model estimations do explain the data findings. In particular, they show that income shocks are transmitted into investment decisions because of the effect of borrowing constraints and income uncertainty, and because the technology of skill formation is highly complementarity in investments across different periods of childhood.

Chapter 4 is based on a work-in-progress co-authored paper with Pedro Carneiro (UCL), Miguel Cordero (U Bristol), Emanuela Galasso (World Bank) and Paula Bedregal (Pontificia Universidad Catolica de Chile). It discusses the use Randomized Control Trial methods combined with dynamic behavioral models of parental investments in children to assess the importance of parental beliefs and attitudes towards child-rearing in parental cognitive and non-cognitive stimulation and several measures of child development. The experimental data is extracted from a large-scale parenting program in Chile aiming at modifying parental beliefs about how to raise their children and their perceived self-competence as parents. The evaluation was designed and implemented by the authors, and we collect pre and post-treatment data from a sample of more than 3,000 households. The investigation has several

innovations: First, it looks at multiple dimensions of parental beliefs, mixing scales borrowed from the literature of psychology of child development and instruments created to elicit parental perceptions, different dimensions of parental investments, and several dimensions of child cognitive and non-cognitive development. Second, we exploit the exogenous variations of beliefs provided by the intervention to estimate treatment effects accounting for measurement error by using a dynamic factor model. Third, we decompose treatment effects by using the theory of mediation analysis to investigate indirect effects of beliefs on child outcomes through changes in parental behavior, and direct changes through changes in the productivity of investments. Finally, we use the experimental data and the elicitation of parental beliefs to separately identify key preference and expectations parameters of a model of parental investments in children in which parents face uncertainty in the returns of the technology of skill formation. Our target is to use the estimated structural parameters to assess the long-term effects in child development of counterfactual policies, with an accent in cost-effectiveness. Preliminary baseline data indicates that parental beliefs are indeed at the root of the socioeconomic gradients we also find in child outcomes and parental investments, and we find some evidence that beliefs potentially impact child outcomes through the two proposed channels. We are currently collecting the follow-up data, and we will be able to estimate treatment effects by October 2014.

The last chapter is a companion chapter which presenting my research agenda in the study of the phenomenon of Labor Informality and Self-employment in developing countries. In this work, I propose a new behavioral model to understand these phenomena of labor markets in developing countries but from a family perspective, in the context where the labor supply, consumption and family formation are all decisions that happen simultaneously. This framework will allow me in the near future to understand gender-specific patterns of labor informality and self-employment both for single and married households, and investigate different potential mechanisms determining the structure of labor markets like comparative advantage, the pension and the welfare system, and intra-household specialization. The final purpose of estimating such a model relies on using the model to perform ex-ante policy simulations of alternative ongoing reforms to labor markets and to the educational system in Latin America, for which there is no evidence of their long-term potential impacts.

# Chapter 2

## Human capital and labor informality

### 2.1 Introduction

The phenomenon of informal labor markets in developing economies has been one of the main concerns for economists and policy-makers during the past decade. According to Gasparini and Tornarolli (2007), nearly 40% of the labor force in Latin America is informal, ranging from lower bounds of 25% in the cases of Chile and Uruguay, to upper bounds of 60% in the cases of Peru and Colombia. The informal sector often comprises small-scale, self-financed and unskilled labor intensive activities, and workers in this sector tend to be younger and less educated (Maloney (1999)). Given that informal activities are fairly widespread across different industries and economic activities, Amaral and Quintin (2006) and Moscoso Boedo and D Erasmo (2013) argue that we can characterize the formal and the informal sectors as having two different production functions. Formal firms not only tend to be larger, but also have larger capital/ labor ratios, higher levels of technology, and demand labor that is more skilled. As a consequence, the formal sector tends to operate at larger ranges of productivity, while wages are also larger (Meghir, Narita, and Robin (2012)). Furthermore, Levy (2010) argues that, in order to avoid costly labor regulations and contributions, informal firms distort their size and tend to operate with suboptimal combinations of capital and labor given the technology available. Since labor and capital are misallocated, observationally similar workers can be considered as less productive in the informal sector. They are also paid lower wages and are not covered by social security. Considering this, there has been a wide consensus that informal labor is composed of workers who, voluntarily or not, do not contribute to the social security system.

Most of the debate on the phenomenon of labor informality has focused on under-



standing whether individuals choose to work informally based on their comparative advantage, or on the contrary, whether this is the result of exclusion driven by segmented labor markets that impose barriers to mobility, in particular towards the formal sector. If markets are competitive, then workers optimally decide on their sector of employment based on their skills and preferences for job attributes like autonomy, independence, or the possibility of avoiding costly taxes and contributions.

In light of this, policies changing labor market incentives in favor of one particular sector would influence the expected rewards of forward-looking individuals and encourage them to accumulate human capital in the sector with the largest expected benefits. Moreover, educational policies facilitating schooling participation could potentially change the incentives for informal labor participation depending on how schooling is rewarded across sectors. In contrast, if barriers to mobility created by poorly designed labor policies prevent workers from choosing jobs according to their skills and preferences there will be little role for policies tackling informality from a human capital accumulation perspective. The evidence from Latin American countries seems rather mixed. While authors like Pagés and Stampini (2007) provide some evidence of labor market segmentation, authors like Perry (2007), Levy (2010), Magnac (1991) and Bosch, Goni, and Maloney (2007) claim that workers self-select into informal jobs based on their comparative advantage.

In this paper, I contribute to this debate by studying, from a dynamic perspective, the extent to which comparative advantage drives participation in informal labor markets. I develop a life-cycle model in which individuals that are heterogeneous in their skills and preferences make decisions about both their schooling and labor market participation into either formal and informal jobs, or non-employment. For this purpose, I explore several mechanisms. First, some authors have noted that in addition to their wages, people might also value some non-wage attributes of a job, which sometimes even compensate them for potentially lower pay (Maloney (2004)). For example, workers who decide to be informal may attach more value to flexibility, autonomy, or the possibility of avoiding paying taxes and contributions, from which they derive little value; while workers in the formal sector may attach more value to the fringe benefits associated with a formal contract. Considering this, I attempt to disentangle the relative importance of wage returns from preferences for sector amenities driving selection into formal and informal jobs, and study whether these valuations vary across education. Second, if the formal and informal sectors can be defined as having different production functions, then individuals will accumulate sector-specific human capital and the returns to different types of

skills should differ across sectors. Moreover, people might have certain unobserved skills that are more productive in one particular sector, making these returns heterogeneous. Consequently, I study whether individuals accumulate different human capital across sectors by breaking down wage differences into sector-specific returns to abilities, schooling, and accumulated experience in several sectors. Third, there is little knowledge of the importance of labor market expectations and dynamics driving both school attendance and labor market participation in the context of an economy with both a formal and an informal sector. That is, if individuals can foresee the existence of two large sectors with potentially different wage returns and job amenities, the assessment of the importance of labor market expectations becomes relevant for the design of policies in the long-term. Finally, I use the model to further study the existence of barriers to mobility across sectors that cannot be explained by skills and preferences, relying on the estimation of transition costs.

Based on these mechanisms, my work contributes to the literature in four different ways. First, I develop a life-cycle structural model building on a Roy model extended to endogenous schooling (Willis and Rosen (1979)), and also extended to compensating wage differentials (Killingsworth (1987)), which I estimate with rich Chilean longitudinal data from a nationally representative sample of households over the period 1980-2009. This approach has important advantages. First of all, it emphasizes dynamics in the decision-making process. Workers may acquire more or less schooling and might change their choice of sector, weighting up current and expected returns from labor markets in dual economies. Additionally, a structural model can shed light on which components of comparative advantage matter more for choices (e.g. abilities, schooling, sector experience, personal preferences). Finally, a structural estimation can also shed light on the presence of barriers to mobility that cannot be explained by skills and preferences by estimating the transition costs of moving across sectors. Addressing these issues using the more traditional approaches previously employed in the literature has several limitations. Some authors have attempted to test for the comparative advantage approach by using wage differentials (Maloney (1999), Yamada (1996)). However, as Magnac (1991) notes, wage differentials might fail to test for selection based on comparative advantage, if workers choose informal jobs based on utility differences associated with non-wage sector attributes. In contrast, a structural approach allows the disentanglement of preferences for non-wage attributes from the observed patterns of choices and wages. Other authors have proposed the study of mobility across sectors as a more reliable test (Bosch, Goni, and Maloney (2007), Bosch and Maloney (2007)). However, as

Pagés and Stampini (2007) argue, in an environment where workers are continuously facing idiosyncratic and industry-specific shocks that require job reallocation, it is still possible to observe mobility in non-competitive labor markets. On the contrary, a structural model enables the simulation of choices within an economy in which workers who have heterogeneous skills and preferences continuously face preference and productivity shocks.

Second, to my knowledge there is little evidence in the literature linking participation in informal labor markets with schooling decisions. Arbex, Galvao, and Gomes (2010) study selection into education and informal jobs in Brazil and find evidence that education is endogenous. However, their approach does not account for dynamics and sector-specific skill accumulation, while these elements are an important source of selection in my findings. In this regard, Pagés and Stampini (2007) study the extent to which education is a passport to accessing better jobs, by analyzing labor market segmentation in three Latin American countries, employing wage differentials and mobility across sectors using longitudinal data. They find evidence of a formal wage premium when these jobs are compared to informal salaried jobs, but no evidence of a formal wage premium when they are compared to self-employment. Nevertheless, in their approach schooling is considered exogenous.

Third, different authors have stressed the importance of heterogeneity when testing for comparative advantage and labor informality. Even when considering the average wage offer to be larger in the formal sector, sector-specific comparative advantage may be driven by heterogeneous skills and tastes; hence for some workers, it is more profitable to be informal. For instance, unobserved skills such as entrepreneurial ability may drive selection into informal jobs associated with self-employment activities. Therefore, I explicitly incorporate permanent unobserved heterogeneity in the form of initial endowments that can be rewarded differently in each sector, with the immediate consequence that wage returns across sectors are heterogeneous. These initial endowments, modeled with discrete and finite unobserved types (Heckman and Singer (1984)), jointly determine school attendance and sector-specific productivity.

I use the structural estimates for returns, preferences and mobility costs to answer two main questions. First, I assess the degree to which comparative advantage drives labor informality relative to market segmentation. I find that human capital accumulation and preferences for job amenities explain up to 72% of transitions between the informal and the formal sector for individuals with less than High School, while labor market segmentation only accounts for 28%. The contributions of com-

parative advantage are decreasing in education (76% for High School degrees, and 83% College). Second, I test for the importance of labor market expectations and persistency in individual decisions. In doing so, I simulate the effect of a recently implemented 20% wage subsidy in formal jobs targeted at workers between 19 and 26 years old, and I find that the subsidy would not only be effective in decreasing informality among the targeted groups, but the incentives to informality also decrease for younger workers (those between 15 to 18 years old). The reduction in informality rates as a consequence of the subsidy would remain persistent until after the age of 40.

Other estimation results are important to be highlighted. First, both wage returns and non-wage job attributes drive choices, but their relative importance varies across education levels. Individuals not completing secondary education value wages more than individuals at higher education levels, who tend to assign a larger relative valuation to non-wage attributes of jobs. Individuals with higher education assign a similar valuation to wages, but tend to value relatively more the associated fringe benefits offered by formal jobs. On the contrary, individuals with low education assign a larger valuation to wage returns in the informal sector. Second, individuals accumulate different types of human capital across formal and informal jobs, and the returns to these skills are heterogeneous. Returns to finish High School are larger in the formal sector, while there is a wage premium for College in the informal sector. Finally, the returns to formal experience are positive in both the formal and informal sectors, while informal experience has positive returns only in informal activities.

The paper is organized as follows: Section 2 provides a short literature review on the theoretical background of informality and its link to the Chilean context; Section 3 describes the surveys and provides the main data descriptives; Section 4 describes the modeling framework; Section 5 discusses the estimation and the identification strategy; Section 6 discusses the estimation results and policy simulations; and Section 7 is the conclusion.

## 2.2 Background and Related Literature

The traditional perspective on why informal labor markets dates from Harris and Todaro (1970) and the ILO (1972)<sup>1</sup>. In this approach, informality is the result of barriers to entry to the formal sector caused by stringent labor regulations like binding minimum wages and segmented labor markets. Magnac (1991) defines labor

---

<sup>1</sup>International Labor Office. 1972. *Incomes, Employment and Equality in Kenya*. Geneva.

market segmentation as a characteristic of dual labor markets in which the rewards in different economic sectors may differ for workers with equal potential productivity and the entry of workers to the formal sector is rationed. One implication of this view is that identical workers will achieve larger benefits in the formal sector, and that they are paid larger-than-equilibrium wages. A second implication is that workers never switch voluntarily from a formal to an informal job.

Recent work has questioned the traditional view of informal work as the disadvantaged sector (Bosch, Goni, and Maloney (2007), Bosch and Maloney (2007), Maloney (1999)). In this literature, workers choose their sector of employment based on vocational choices and their comparative advantage to work in a more entrepreneurial sector. Thus, informality status may be driven by choice rather than exclusion (Perry (2007)). Under this view, labor markets are competitive, there are no barriers to mobility across sectors, and the formal and the informal sectors are symmetrical, equally desirable, and competitive with different production functions. Levy (2010) proposes a more nuanced view of the comparative advantage approach, recognizing that labor markets are not necessarily competitive and that costly labor regulations cause some distortions. He argues that the informal sector is less productive than the formal sector because there is a misallocation of capital and labor across the sectors produced by badly designed social policies, such as social protection programs for the poor, which induce a higher than optimal rate of firms and workers operating informally. On the one hand, firms optimize given the constraints imposed by labor regulations, so in order to operate formally they pay higher labor costs, are more productive and have better technology. Complementarity of skills and technology means that in equilibrium formal firms demand more skilled workers. On the other hand, workers choose employment optimally but this is constrained by their skills and their tastes for non-wage sector attributes like autonomy, flexibility, or the possibility of evading taxes and social security contributions (Maloney (2004)).<sup>2</sup>

Most of the evidence in Latin American countries supports the comparative advantage approach to informality. Analyzing wage differentials, Maloney (1999) for Mexico and Yamada (1996) for Perú find little evidence that formal workers have

---

<sup>2</sup>Moscoso Boedo and D Erasmo (2013) develop a macroeconomic model of TFP with capital imperfections calibrated with data from developing countries, and find that countries with a low degree of debt enforcement and high costs of formalization are characterized by low allocative efficiency and a larger informal sector, lower productivity, and lower stocks of skilled workers. Paula and Scheinkman (2007) develop and test an equilibrium in which they show that managers in formal firms have higher levels of ability and choose a higher capital-labor ratio than informal entrepreneurs.

higher earnings than the self-employed. Moreover, they find evidence of positive selection into micro-entrepreneurial activities. But as Magnac (1991) notes, testing for competitive labor markets by comparing observed or potential wages is incorrect because of selection bias; workers might have specific skills in each sector. Instead, he tests for segmentation using data for females in Colombia by incorporating entry costs into a standard Roy Model, and finds that the assumption of competitive labor markets cannot be rejected. One limitation of his approach is that if utility depends on the non-wage-related attributes of jobs in each sector, segmentation is no longer testable using a Roy model but should be tested using a compensating wage differentials model. Pagés and Madrigal (2008) use job satisfaction data from three low-income countries to assess the extent to which different types of informal jobs provide compensating amenities. They find a large degree of heterogeneity of job valuation within informal jobs and across formal and informal jobs. For example, within typically classified informal jobs, self-employment activities are the most preferred, while being an informal salaried worker in a small firm is the least preferred.

The limitations of the analysis of wage differentials have led some authors to test for the comparative advantage approach by studying mobility across sectors. Bosch, Goni, and Maloney (2007), Bosch and Maloney (2007) and Maloney (1999) find substantive evidence of mobility across the formal and the informal sectors, with higher rates among the less-skilled. Bosch, Goni, and Maloney (2007) argue that a substantial amount of the informal work corresponds to voluntary entry, which is particularly true for the self-employed. Nonetheless, they also recognize that informal salaried work may correspond closely to the standard queuing view, especially for younger workers. Pagés and Stampini (2007) obtain similar results by developing a benchmark mobility indicator measuring the degree of mobility that would occur in a world in which all states are equally preferred and compare the actual rates of mobility to that indicator to test for segmentation. They find evidence of labor market segmentation when comparing the formal salaried to the informal salaried jobs, whereas self-employment participation is driven by comparative advantage.

Finally, some authors have provided evidence that supports the comparative advantage approach to informality by investigating the returns to schooling across sectors. Amaral and Quintin (2006) and Paula and Scheinkman (2007) find heterogeneous returns to college education in the informal sector in Argentina and Brazil. Arbex, Galvao, and Gomes (2010) note that in order to work informally, skilled workers have to give up some fringe benefits associated with a formal con-

tract. Therefore, returns to college education should be positive or at least high enough to offset the lack of benefits. Developing a two-period theoretical framework with endogenous schooling and heterogeneous returns, tested empirically by using IV quantile regressions with Brazilian data, they find an education premium in the informal sector, which varies along the conditional distribution of earnings. Meghir, Narita, and Robin (2012) develop an equilibrium search model with a formal sector and an informal sector in Brazilian labor markets, and find that on average wages in the formal sector are higher than in the informal sector. However, informal workers are paid more than formal workers in firms operating at the same level of productivity.

Some important implications can be extracted from the available evidence for modeling considerations. First, to overcome the limitations of the analysis of wage differentials, I propose a structural estimation that considers self-selection into informal jobs and self-employment based on both wage differentials and non-wage sector amenities, which might be valued differently for workers at different education levels. Second, I explicitly model transition costs in order to capture persistency in choices found in the data, and to test for the existence of additional sources of barriers to mobility which cannot be explained by skills and tastes. And finally, the incorporation of heterogeneous returns is a key factor to capture all the different dimensions of comparative advantage that have been previously discussed in the literature.

## 2.3 Data and institutional framework

### Institutional framework

I consider the informal sector as being composed of firms that are not registered with the authority, not paying taxes, and not either paying social security contributions or coming under the labor laws; and by all full-time (more than 40 hours a week) salaried and self-employed workers reporting that they are not covered by social security contributions.

The evidence of labor market segmentation in Chile is rather scarce. Contreras, Mello, and Puentes (2008) argue that Chile's tax system is not particularly burdensome, and with regard to labor regulations, the Chilean dictatorship during the 1980s strongly deregulated the labor markets, decreasing severance pay, dismissal costs and minimum wages, and prohibiting union activity. A reform in 1980 intended to link contributions with benefits transformed the pay-as-you-go social security system into a full capitalization system, including pensions and health in-

insurance, making Chile the least regulated labor-market in Latin America. Heckman and Pagés (2003) argue that the incentives for informality from social protection programs for the poor are very small in Chile, compared to bigger economies like Mexico, Brazil or Argentina.

Further characteristics make the Chilean labor market attractive for the study of labor informality using a comparative advantage approach. First, social security contributions are voluntary for the self-employed. As a consequence, the large majority of self-employed workers are informal, particularly those with lower levels of education. Second, social security contributions and taxes are compulsory for employees, and employers are responsible for deducting them automatically from their salaries. However, the labor protection rules can be easily avoided by small firms, and as a result roughly half of the informal workers are salaried employees (the other half are self-employed). Finally, Contreras, Mello, and Puentes (2008) provide some evidence that the more educated tend to hold more formal jobs, and that participation is the result of self-selection based on skills rather than the effects of barriers to mobility, which is tested in the context of a wage differentials approach.

With regard to educational institutions, two aspects are important to highlight. First, due to massive liberalization of the education market in 1981, private provision of schooling at both high school and college levels is very high. At tertiary level, average tuition fees are very high and show high variability (US\$2,700 a year compared to an annual minimum wage of US\$ 3,800). As monetary costs for schooling are large and might greatly influence schooling decisions, I incorporate the variability in the data on tuition fees in order to identify college choices in the modeling framework. Tuition fees at secondary level are rather low. In the Chilean school system three schooling systems co-exist: free public schools (50%), private subsidized schools (43%) and private non-subsidized schools (7%). The amount of subsidies received by the second group are as large as the cost per student that the state spends on public schools, so the variability in the tuition fees paid by the families in the sample is rather low to be used as a source of identification of high school choices. For this reason, I do not include monetary costs in the preferences for high school participation.

## The surveys

The “Encuesta de Protección Social” (Social Protection Survey) is a longitudinal survey containing four waves: 2002, 2004, 2006 and 2009. It covers a nationally



representative sample of 14,045 individuals who are followed across the four waves with very low attrition rates. In the first wave, individuals are asked to report their family background, all of their educational history, and all of their labor activities from 1980 onwards, which include the type of job performed, hours of work, whether they were paying social security contributions in that job, and their labor status (whether they worked in a firm or were self-employed). Since female labor participation is rather low (44%), the model is estimated for males to avoid fertility decisions.

Direct costs to schooling like tuition fees, are not observed in the sample. In order to simulate college choices I use a second data source, the CASEN survey, to construct a tuition fee index by municipality and year, which is incorporated as monetary costs into the model. This survey is nationally representative, and among many other socio-demographic variables, households report the fees they were entitled to pay, and any amount of subsidy provided by the state, so the total monetary costs can be retrieved. The data is available for the years 2000, 2003, 2006 and 2009, coinciding with the panel of wages and choices used for estimation. In order to control for potential sources of endogeneity of concurrent trends of tuition fees and labor outcomes, the data is time and municipality detrended before being used for the simulations.

Finally, an important drawback of the panel survey is that information on wages is only available from 2001 onwards. As high school and college participation sharply increased during the 1980s and the first half of the 1990s in Chile, returns to schooling are not expected to be the same across cohorts. To overcome this data limitation, I reconstruct the wage profiles by schooling for the oldest cohorts at younger ages (for which I do not observe wages), by assuming that cohort effects by schooling are constant across ages. Furthermore, the model is estimated using the data for individuals who started making choices after 1980, as it's not possible to track sector experience in each sector for those who began their labor market participation earlier.

In total, the panel of males consists of 4,493 individuals, with 117,003 individual-year observations.

## Data descriptives

Some data descriptives are worth showing to shed light on the main correlations and sources of dynamics.

Figure 2.1 describes informality rates for the males by age group over the period

of the study. Informal work is more common among the youth and the elderly, trends which remain after controlling for cohort fixed effects <sup>3</sup>. The u-shape of informality rates is explained by a larger amount of informal salaried work among the youth, and increasing participation in self-employment when workers become older. Figure 2.2 also shows that informal labor participation decreases for more educated workers, a trend which remains relatively stable over the life-cycle. These trends are similar for women<sup>4</sup>. Remarkably, the strongest differences in informality arise between high school degrees and lower than high school levels. Schooling differentials remain when the informality rates are analyzed separately for the self-employed and salaried employees. <sup>5</sup>.

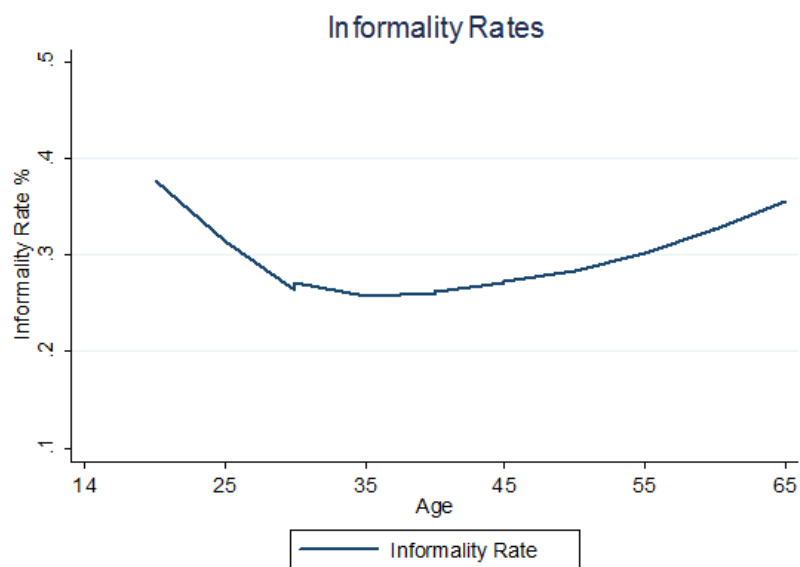


Figure 2.1: Mean Informality Rates Males

---

<sup>3</sup>(Figure A.1 Appendix A)

<sup>4</sup>(Figure A.2 Appendix A)

<sup>5</sup>(Figure A.3 Appendix A)

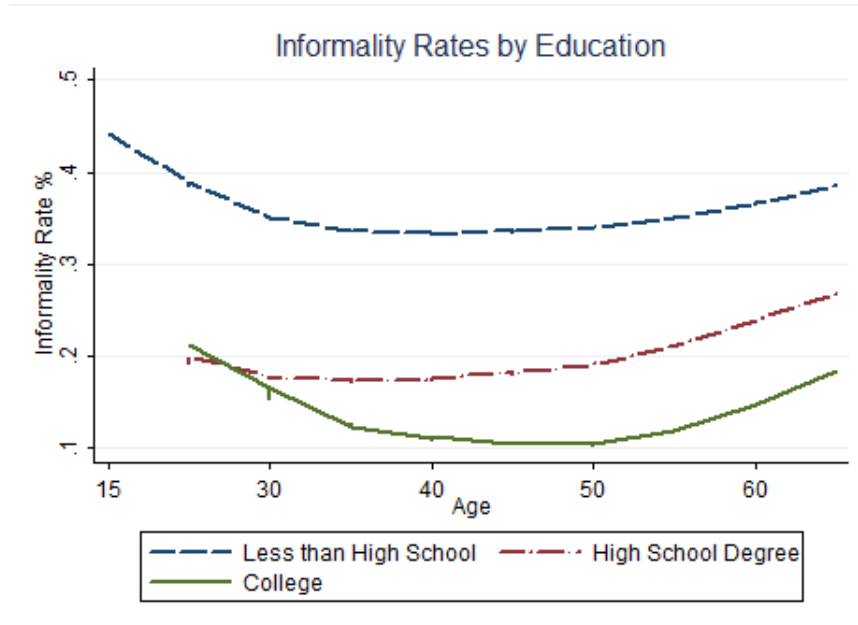


Figure 2.2: Informality Rates by Education

Figure 2.3 compares net wages (after tax and social contributions) in a two-sector economy. Consistent with evidence from other Latin American countries, formal wages are always larger than their informal counterparts for all schooling levels. However, the formal wage premium is neither constant over the life-cycle nor across schooling levels. For high school dropouts, the average wage premium is relatively stable for all age groups, whereas for individuals with high school degrees and college level education the wage premium widens. This feature, and the fact that wage profiles have different slopes across education groups, suggest different underlying dynamics of sector-specific human capital accumulation, justifying the choice of a comparative advantage approach as the modeling framework<sup>6</sup>.

<sup>6</sup>Note that wage profiles with different slopes across education groups require the inclusion of interactions between education and sector experience in the estimation of standard Mincerian wage equations.

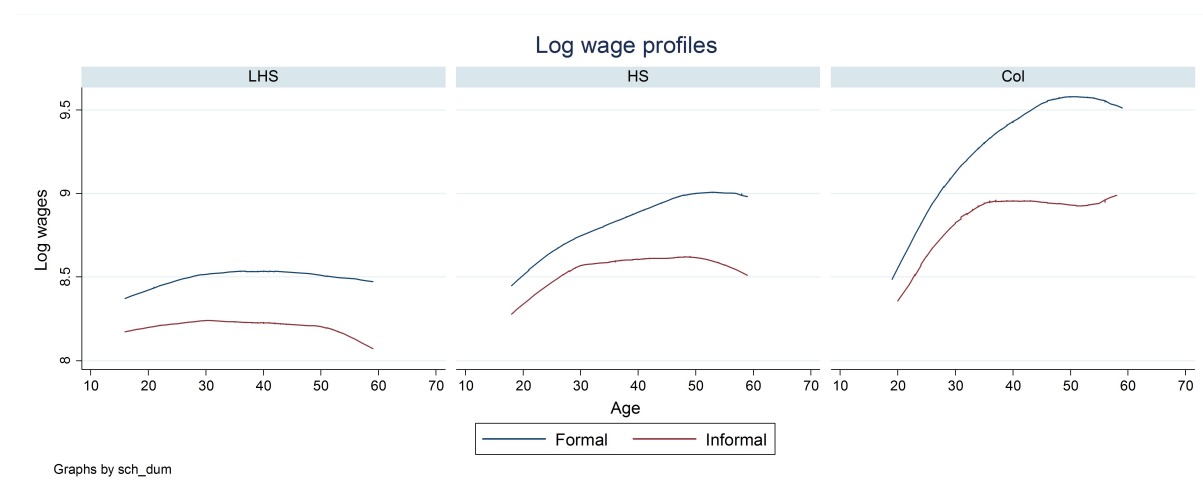


Figure 2.3: Mean wages by education in two sectors: formal and informal employment

As discussed above, Maloney (1999) and Pagés and Madrigal (2008) indicate that there is evidence of heterogeneity within informal jobs. The self-employed tend to report higher levels of job satisfaction compared to the informal salaried, and they are likely to self-select into these jobs while the informal salaried seem to face some barriers to mobility. Figure 2.4 describes participation patterns in Chilean labor markets. While informality among the salaried decreases over the life-cycle, self-employment increases. As roughly 85% of the self-employed do not contribute to the pension system, and therefore are informal, the patterns of informality described in Figure 1.a are the result of composition effects, which would be important to the interpretation of my results. Youth informal workers are mainly salaried workers employed in small firms, whereas the high number of informal workers among the elderly reflects a larger proportion of self-employed. Indeed, Perry (2007) argue that some workers with entrepreneurial abilities start their working lives as salaried employees where they accumulate capital and experience to run their own businesses later in life. These patterns are consistent with previous evidence that informal salaried work is less desirable than self-employment.

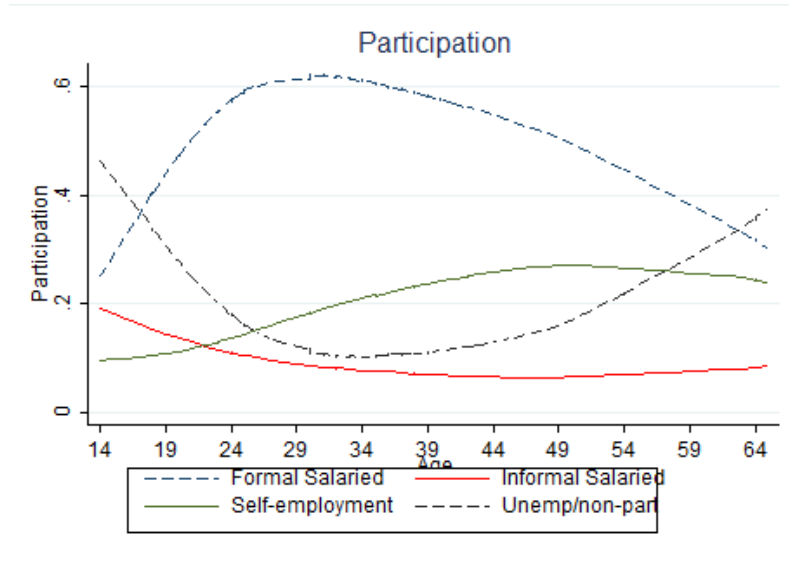


Figure 2.4: Labor market participation by sector

Table 2.1 indicates that the probability of having some level of experience as informal or as a self-employed in the sample is rather large, so there is mobility across sectors. The fraction of workers with experience as both formal and informal, and as self-employed workers, decreases with schooling, which is consistent with the data for other Latin American countries reported by Perry (2007). On the one hand, 69% of high school dropouts have labor experience both in the formal and in the informal sectors, while these rates are lower for individuals with high school degree and college level (52% and 40%). On the other hand, 70% of high school dropouts have experience as salaried employees and of being self-employed, while these rates also decrease for individuals with high school degree and college level (32% and 13%).

Schooling Level	Experience as formal and informal	Experience as self-employed and salaried employee
LHS	69%	70%
HS	52%	32%
Col	40%	13%

Table 2.1: Fraction of workers with experience as formal/informal, and self-employed/salaried employees

However, once workers enter a sector, the probability of switching to another sector is rather low, or persistency in choices is large. For example, among the formal salaried workers, the probability of remaining in the same sector is more than 90%. In contrast, transition rates to the formal salaried state conditional on being informal salaried are larger, and they tend to increase for more educated workers (9.5% for LHS, 16.6% for HS and 15.7% for College). This seems to be in line with the idea that informal salaried work responds to the traditional view of barriers to entry to formal jobs. Conditional on being self-employed, the probability of moving to the formal salaried sector is 5% in average (4.2% for LHS, 5.9% for HS and 5.4% for College), roughly twice the probability of moving to self-employment from the formal salaried state. Finally, note that the degree of persistency in the self-employment state is almost as large as in the formal salaried state, and much larger than persistency in informal salaried jobs. This is consistent with the notion that self-employment is a choice (Bosch, Goni, and Maloney (2007)).

	Less than High School				High School Degree				College			
	F	I	S	U	F	I	S	U	F	I	S	U
F	90.1%	1.8%	2.5%	3.7%	90.7%	1.1%	2.1%	2.3%	90.8%	0.6%	1.7%	2.4%
I	9.5%	80.2%	3.7%	5.6%	16.6%	67.6%	4.9%	5.0%	15.7%	61.0%	3.9%	10.1%
S	4.2%	1.5%	89.6%	3.2%	5.9%	1.2%	84.9%	2.1%	5.4%	1.2%	83.5%	2.1%
U	12.1%	4.1%	4.3%	77.4%	20.5%	4.3%	3.8%	65.1%	18.6%	3.7%	3.9%	67.2%

Table 2.2: Transition probabilities across jobs from period  $t$  to  $t + 1$

## 2.4 The Model

The theoretical framework is an extended Roy Model allowing for endogenous schooling choices (Willis and Rosen (1979)) and for compensating wage differentials (Killingsworth (1987)). I build on the literature on career paths and dynamic discrete choice models with unobserved heterogeneity proposed by Keane and Wolpin (1997) and Adda, Dustmann, Meghir, and Robin (2013). Two important assumptions are considered for modeling purposes. The first assumption is a partial equilibrium environment, so I analyze workers' decisions given the incentives provided by the current structure of labor markets; therefore the skill rental prices are fixed and known to the individual. Under the presence of GE effects, the skill rental prices can change with the relative supply of workers with different types of skills, and individuals internalize the change in returns in their decisions. Therefore, the analysis of how people react to education and labor market incentives should be taken as middle-term responses.

The second assumption in this particular setup is risk neutrality. The inclusion of risk aversion has a potential effect on labor supply in two cases. First, individuals might choose to be more formal or informal and save for precautionary reasons in response to potentially larger unemployment or income shocks inherent to one particular sector, or in the case of informal work, in response to the uncertainty of not having a pension on retirement. However, it has been reported that in Chile savings are zero or negative for the first four income quantiles<sup>7</sup>. And second, risk averse workers might react to the incentives provided by the contributory pension account, which are compulsory savings for formal workers. But as Attanasio, Meghir, and Otero (2011b) note, these incentives are likely to affect labor supply decisions only for workers who are more than 45 years old, and they report that for males in these age groups the effects are very low (less than 1.7% decrease in formal labor participation). Nonetheless, an interesting extension of the model is the incorporation of risk averse individuals alongside both private and pension savings to analyze how schooling and labor supply would change in response to credit market imperfections.

In this modeling framework, workers then choose schooling and sector participation according to their comparative advantage, but transitions across sectors are costly, which might reflect the presence of labor market rigidities. A worker's comparative advantage is a complex vector, which includes observed and unobserved skills, and tastes for non-wage job amenities. Unobserved skills are modeled as initial skill endowments by including discrete and finite unobserved types in the fashion of Heckman and Singer (1984). Wage offers are sector-specific and are the realization of a technology of skill production function that embodies the accumulated human capital of an individual valued in a particular sector according to an equilibrium rental price. The model attempts to reproduce the dynamics found in the data by explicitly incorporating labor market expectations as part of the valuation of current choices. In the context of a human capital investment model, this means two things. First, past schooling and labor choices determine the accumulated level of skills, which could be rewarded differently in each sector depending on market prices. And second, individuals don't know for sure the future benefits of a particular choice in the present, but they know the distribution of shocks, so they can evaluate the future expected rewards for every possible current choice.

It is worth stressing some model mechanisms. First, wage returns are heterogeneous. Therefore, the fact that formal wages are larger on average doesn't mean that this is true for everyone. Furthermore, conditional on schooling and sector-

---

<sup>7</sup>Butelmann and Gallego (2000). "Household Saving in Chile: Microeconomic Evidence". Central Bank of Chile.

experience some individuals might earn more in informal salaried jobs or in self-employment if their initial skill endowments are better rewarded in these sectors, which reflects the fact that skill endowments are not necessarily equally productive across sectors. Second, initial skill endowments explicitly relate to schooling choices and productivity across sectors, as skill endowments act as a proxy for the underlying ability of the individual, which might determine self-selection into higher levels of schooling. Third, even if the accumulated human capital of an individual is better rewarded in a particular sector, she might end up choosing another sector for several reasons. For example, she might value the non-wage attributes of the job more, like the level of flexibility or autonomy, or she might face large search or psychological costs to transition to the sector with a larger wage. Finally, these features have important implications for understanding how people would react to incentives. For example, the effects of individual behavior of a decrease in income taxes in the formal sector are reduced if non-monetary incentives are too important or if schooling attainment or the choice of a sector are highly dependent on ability.

The timing of the model is as follows. Individuals make their first choice at age 14. They can achieve three education levels: “Less than High School”, “High School Degree” or “College”. Everyone starts with primary schooling which is completed at  $t = 0$  (age 14). In the sample 96% of students complete this level. At every subsequent period people decide whether to continue to the next schooling level, start working in formal or informal jobs, or stay out of the labor market in unemployment or home production activities. Denote  $m = \{F, I\}$  one of the three working sector choices, where  $F$  =formal salaried and  $I$  =informal salaried. The choice of home production, non-participation or unemployment is denoted by  $U$ .

### 2.4.1 The State Space

Denote the state space  $\Omega$  as the set of variables that define the state-dependency of individual utilities over time. I estimate a life-cycle model so the time dimension  $t$  is the age of the individual. I detrend the data on wages and tuition fees to control for macroeconomic trends, and I take out cohort effects to compute data moments.  $Ed_{it}$  is the schooling level of individual  $i$  at age  $t$ . Then  $Ed_{it} = \{LHS, HS, Col\}$  or Less than High School, High School Degree level, and College level. Accordingly, labor experience accumulates by sector.

Finally, I denote the unobserved ability by  $\mu_i$ . Higher ability individuals self-select more into schooling, and at the same time, people have different sets of skills that make them more productive in one sector than in another, driving their choices.



For example, entrepreneurial ability might drive self-selection into informal jobs or self-employment, while the ability to work in structured work environments might drive self-selection into formal jobs. Both of these are known by the individual and fixed from age 14, or  $t = 0$ . On the other hand, they are unobserved by the econometrician and need to be estimated along with the rest of the parameters. I incorporate permanent unobserved heterogeneity by modeling a discrete number  $k$  of unobserved types (Heckman and Singer (1984)), so that  $\mu_k$  is an indicator variable for type  $k$ . The use of a discrete and finite number of types is important to make the dynamic programming problem tractable.

### 2.4.2 Flow Utilities

At every period, individuals derive an instantaneous utility from attending school, staying at home or working in an economic sector. The costs of that decision are, in the case of schooling, foregone expected earnings or rewards from leisure/home production. When choosing the working sector, I also allow for non-monetary rewards coming from sector-specific job attributes.

Denote the vector of available choices at  $t$  by  $\{Ed, m, U\}$ . Since everyone at  $t = 0$  starts with  $Ed_{it} = LHS$ , people can make further schooling choices only for High school degree or College, then  $Ed = \{HS, Col\}$ . Denote  $U_{it}^{Ed}$  the instantaneous utility of attending schooling at level  $Ed$ . Then,

$$U_{k,R}^{Ed} = \gamma_{1,k}^{Ed} \mu_k - TC_R^{Ed} - \eta_t^{Ed}$$

where  $TC_{iR}^{Ed}$  is the tuition fee index constructed from household data, which varies by schooling level  $Ed$  and municipality  $R$ . The index is constructed with detrended costs varying over time and across municipalities in order to control for potential concurrent trends with wages and choices. The factor load  $\gamma_{1,k}^{Ed}$  represents psychic costs or the consumption value of the schooling decision, which may capture both heterogeneous abilities and family background which translates into financial constraints to attending high school or college. The term  $\eta_t^{Ed}$  is a preference shock to schooling including the non-monetary costs of school attendance not observed in the data.

The utility of working in sector  $m$  is

$$U_{t,k}^m = \gamma_{2,Ed}^m W_{t,k}^m - \gamma_{3,Ed}^m + \gamma_4^m t + \gamma_5^m t^2$$

where  $W_{tk}^m$  is the gross wage offer the individual type  $k$  observes at age  $t$  in sector  $m$  and  $d_{it}$  is the choice the individual makes at age  $t$ .<sup>8</sup> Several additional terms affect sector preferences:  $\gamma_{2,Ed}^m$  is the wage valuation, which emphasizes both the marginal utility of income and the valuation of some sector amenities that are non-separable in the utility function. Additional sources of heterogeneity are included in this parameter by allowing it to vary across different sectors and education levels.  $\gamma_{3,Ed}^m$  captures the worker's valuation of non-wage sector amenities which are separable in the utility function. One can interpret this parameter as fixed costs of working, which are also allowed to vary by education levels. This parameter is relevant to understand by how much individuals compensate wage differentials with non-wage sector attributes. A quadratic function in age is also incorporated to capture the fact that individuals at different ages might have different tastes for insurance and/or labor market participation in one particular sector. The parameter capturing age effects is  $\gamma_4^m$  and  $\gamma_5^m$ .

The utility of unemployment/leisure/home production is

$$U_{t,k}^U = \gamma_{1,k}^U \mu_k - \eta_t^U$$

The reward that individual type  $k$  obtains from staying at home depends on unobserved skills captured by  $\gamma_{1k}^U$ , and preference shocks  $\eta_t^U$ , which is a random component reflecting uncertainty in the valuation of leisure or home production. For example, pregnancy can increase the valuation of unemployment for women.

Finally, the model also explicitly includes transition costs, which are explained below in the section on value functions. These costs are required to be included in order to match the high levels of persistence found in the data.

---

<sup>8</sup>In a model with risk neutral individuals and returns from the stock markets similar to the interest rate, the timing does not matter when social security contributions such as pension retirement are taken into account. To the extent that there is full capitalization, as in the Chilean case, I consider gross wages, which in the formal sector already account for the total compensation including social security contributions.

### 2.4.3 The Wage Offer

Every time individuals choose to study they forego earnings from work. Since we have different working sectors, every time individuals choose a sector they also forego earnings in another sector. Wages are a function of the skill production function  $H_{t,k}^m$  and skill rental prices  $r_t^m$ . As it was specified above, skill functions vary by sector reflect the existence of different production functions across the formal and the informal sector, which might translate into different marginal productivities of each skill component. The human capital accumulated by workers is a function of their schooling level, their accumulated sector experience in the same sector and across sectors, sector-specific unobserved abilities, and productivity shocks.

$$W_{t,k}^m = r_t^m H_{t,k}^m = r_t^m f(Ed, X^F, X^I, \mu_k, \epsilon_t^m)$$

Given a functional form for the skill function (exponential in this case), the sector-specific log wage offer can be defined as follows,

$$\ln W_{t,k}^m = \alpha_{0,k}^m \mu_k + \alpha_1^m HS + \alpha_2^m Col + \alpha_{3,Ed}^m \ln(1 + X_t^F) + \alpha_{4,Ed}^m \ln(1 + X_t^I) + \epsilon_t^m$$

where  $\alpha_{0,k}^m$  represents the rental price for the initial endowment in sector  $m$  for individual type  $k$ , and captures selection across choices.  $\alpha_1^m$  and  $\alpha_2^m$  capture the average returns to schooling levels. Since the data wage profiles by sector and education show very different slopes, which also vary along the life-cycle, I estimate returns to experience varying by education level, captured by the parameters  $\alpha_{3,Ed}^m$  and  $\alpha_{4,Ed}^m$ .

Finally, in order to persistency in wages from unobserved factors, I include persistent productivity shocks whose sector-specific innovations are allowed to be correlated across sectors

$$\epsilon_t^m = \rho^m \epsilon_{t-1}^m + \xi_t^m$$

$$\xi_t^m \sim N(0, \Sigma)$$

### 2.4.4 Uncertainty

The source of uncertainty in the model comes from preference and productivity shocks. Preference shocks are modeled following an extreme value type 1 distribu-

tion, while innovations in persistent wage shocks are normally distributed. Shocks are important to produce mobility across all choices in  $t$  because they shape expected utilities in each of the alternative choices from  $t + 1$  to  $T$ , affecting rewards from current choices. It is likely that innovations of productivity shocks are correlated across choices; that is why I draw wage shocks from a multivariate normal distribution. Sector-specific autocorrelations and the distribution of innovations are identified by the time series properties of wage data, so these parameters are estimated along the rest of structural parameters.

### 2.4.5 Recursive problem

Self-selection into schooling and jobs is based on expected earnings, which depend on current choices. This entails non-separability over time. The model dynamics are as follows: all individuals start at age 14 having finished primary school. This is a fairly safe assumption as 96% of the sample actually did finish primary schooling (in Chile primary school has effectively been compulsory since 1962). Every year they must choose whether to continue studying for an extra year of secondary schooling, dropout and start working or, stay out of the labor force. At age 18 they must decide whether to continue to College level or drop out of education. If they drop out of school before the 4th level of high school then their education level stays at  $Ed_t = LHS$  (Less than High School). If they dropout straight after the 4th level of secondary school then  $Ed_t = HS$  (High School degree), and if they continue studying to College level then  $Ed_t = Col$  (College). The maximum level of schooling is standardized to 5 years of College. If an individual drops out at any schooling level, she cannot go back to school, a fact that is supported by the data.

If an individual decides to switch sector, she pays a fixed cost  $c^{i,j}$  to move from sector  $i$  in  $t - 1$  to  $j$  in  $t$ . The purpose of these costs is to capture high levels of choice persistence in the data, which can be driven by several factors like search costs, skill depreciation or psychological costs of transitions. They might also capture certain labor rigidities that could partially explain mobility rates.

Denote by  $\Omega_t = \{Ed_t, X_t^F, X_t^I, \mu_k, R, \epsilon_t^m\}$  the state space of type  $k$  at age  $t$ .  $R$  is an additional state variable saving the municipality where the individual lived while studying and is used to assign direct costs to schooling. This implies that in the first two periods the model the Value Function is solved for each of the 300 municipalities in the sample. Then the value of education at level  $Ed = \{HS, Col\}$  is

$$V_{t,k}^{Ed}(\Omega_t) = U_{t,k}^{Ed} + \beta Emax \left\{ V_{t+1,k}^{Ed}(\Omega_{t+1}), V_{t+1,k}^m(\Omega_{t+1}) - c^{Ed,m}, V_{t+1,k}^U(\Omega_{t+1}) \right\}$$

By choosing schooling individuals obtain the instantaneous utility  $U_{it}^{Ed}$  plus the discounted expected maximum value over the available alternatives at  $t+1$ : continuing to the following schooling level, working in one of the two sectors  $m = \{F, I\}$ , or staying out of the labor market. Expectations are taken over the distribution of preference and productivity shocks implied by choices. Notice that  $X_{it+1} = X_{it}$  when individuals choose schooling and that they pay a transition cost only when moving to work, but not when moving to unemployment/leisure/home production.

Similarly, the value of working as a formal-employee at  $t$  is,

$$V_{t,k}^m(\Omega_t) = U_{t,k}^m + \beta Emax \left\{ V_{t+1,k}^m(\Omega_{t+1}), V_{t+1,k}^{m'}(\Omega_{t+1}) - c^{m,m'}, V_{t+1,k}^U(\Omega_{t+1}) \right\}$$

where it is clear that the worker cannot go back to school, and if she wants to switch sector, she has to pay a transition cost  $c^{m,m'}$ . By choosing sector  $m$ , individuals accumulate another year of experience in that sector, which translates into an increase in the valuation of all choices rewarding labor experience in sector  $m$  at age  $t+1$ . Finally, the value of unemployment/leisure/home production is

$$V_{t,k}^U(\Omega_t) = U_{t,k}^U + \beta Emax \left\{ V_{t+1,k}^m(\Omega_{t+1}) - c^{U,m}, V_{t+1,k}^{m'}(\Omega_{t+1}) - c^{U,m'}, V_{t+1,k}^U(\Omega_{t+1}) \right\}$$

where the choice of unemployment does not alter the state space for the next period.

### 2.4.6 Mobility

In the model, mobility across sectors is generated by three sources. First, an individual may switch to another sector if there is a large positive shock in the sector she intends to move to, and the gains in productivity due to this shock and to the returns to skills in the new sector are larger than the mobility costs and the potential losses in the returns to skills if she stays. Second, even if the shocks and transition costs across sectors are exactly the same, the worker might still want to switch if an additional year of experience in the new sector is better rewarded than an additional year of experience in the current sector. Note that experience accumulated in each sector is potentially rewarded in all sectors with different rental prices. Finally, mobility costs across sectors might prevent individuals from switching. For example, if at some point in the life-cycle a low-skilled informal worker faces a negative shock,

she might consider switching to a similar formal job as her experience would also be rewarded in the new job. However, this decision might be prevented by unaffordable entry costs, search costs, rationing, or the lack of networks, preventing movement.

## 2.5 Model Solution and Estimation

### 2.5.1 Solution Method

Dynamic discrete choice models do not have an analytical solution. Within a finite horizon context, the model must be solved numerically using backward recursion methods. At period  $T$ , each individual draws random shocks from the multidimensional error vector  $(\eta_T, \epsilon_T)$  and chooses the alternative that yields the maximum instantaneous utility evaluated at every possible state space combination of schooling and labor histories. I assume that the terminal value function over the life-cycle is  $V_{T+1} = 0$ . I denote  $d_t^* = \{Ed, m, U\}$  the optimal choice at every period. Then, at period  $T$  individuals solve

$$d_T^* = \operatorname{argmax}(U_T^{Ed}, U_T^m, U_T^U)$$

At period every period  $t$ , two steps are required to compute the value functions. First, they need to evaluate expectations over  $t + 1$  computing the  $E_{\max}$  functions, where expectations are taken over  $(\eta_{t+1}, \epsilon_{t+1})$ , evaluated at every possible choice and state space combination at  $t$ .

To solve for the fact that wage shocks in  $t + 1$  depend on the realizations of wage shocks in  $t$ , I follow Galindev and Lkhagvasuren (2010) to approximate persistent shocks in more than one dimension with innovations which are potentially correlated across dimensions. They adapt Tauchen's method (Tauchen (1986)) by using Markov Chain processes, which they prove is an efficient method provided that the autocorrelation parameters are not close to unit root<sup>9</sup>. The evaluation of the  $E_{\max}$  function then involves a multidimensional numerical integration as follows,

$$E_{\max}[V_t^{Ed}, V_t^m, V_{Tt}^U] = \int_{\epsilon_t} \left\{ \int_{\eta_t} \max[V_t^{Ed}, V_t^m, V_t^U / d_{t-1}^*, \Omega_{t-1}, \epsilon_{t-1}] f(\eta) d\eta \right\} f(\epsilon_t | \epsilon_{t-1}) d\epsilon$$

<sup>9</sup>The analysis of the time series of the wage residuals show that the autocorrelation parameters in both sectors are close to 0.9

Where  $f(\epsilon_t | \epsilon_{t-1})$  is the transition matrix for the Markov process of wage shocks. This matrix is a function of the autocorrelation parameters  $\rho^m$  and the variance of wage innovations  $\Sigma$ . The advantage of modeling preference shocks  $\eta_{it}$  with an extreme value Type I distribution is that the expected value (*Emax*) has a closed form expression so we can decrease the dimensions of numerical integration. Therefore, the evaluation of the *Emax* function only involves the numerical integration across the dimensions of wage shocks, which are normally distributed.

Second, I evaluate the instantaneous utilities at  $t - 1$ , again for every possible combination of the steady state at that period, drawing the error vectors  $(\eta_{t-1}, \xi_{t-1})$  and compute the value functions at  $t - 1$ :  $(V_{t-1}^{Ed}, V_{t-1}^m, V_{t-1}^U)$ . The optimal choice at  $t - 1$  is then obtained from

$$d_{t-1}^* = \operatorname{argmax}(V_{t-1}^{Ed}, V_{t-1}^m, V_{t-1}^U)$$

The process is then repeated in the same fashion until  $t = 0$ , where the outcome is the evaluation of the optimal choice  $d_t^*$  for every combination of the state space  $\Omega_t$  in every period.

### 2.5.2 Model Identification

Three aspects of model specification are worth discussing in my modeling framework. First, the model does not suffer from an initial condition problem. As Aguirregabiria and Mira (2010) note, in a model with unobserved heterogeneity, if the initial state space varies across individuals in the sample, one needs to fully specify how the distribution of unobserved heterogeneity changes with initial states. Initial endowments in the first period of the model are likely to be correlated with observable states. Therefore, if there is variation in the distribution of initial states in the sample, one would need to use some parametric or non-parametric specification of how individuals made choices in the past conditional on unobservables, and solve it backwards until the state space becomes independent of permanent unobserved heterogeneity. In this particular case, almost everyone in the sample finished primary school (primary schooling has been compulsory in Chile since 1962), and thus everybody started with the same education and experience.

To identify wage returns in different sectors, I exploit a large sample variation of wages by sector, schooling, and sector experience, to estimate counterfactual wage

returns. I observe data for individuals from different cohorts, so I also exploit time variation to reconstruct wage and participation moments by age. I control for cohort effects in sector wages to incorporate the potential variation in returns over time as a result of merging data from different cohorts. With regard to the identification of unobserved heterogeneity parameters varying by type, I attempt to identify the whole distribution of wage profiles by matching different percentiles of sector wage distributions by schooling.

In order to identify the preference parameters in non-working activities, Todd and Wolpin (2010) note that one only needs to specify an exclusion restriction in the working alternatives. In this case, observable education, sector experience, and accepted wages are sufficient statistics to identify those parameters. The identification of preference parameters in the working alternatives require further exclusion restrictions in the non-working alternatives. I exploit the variability on tuition costs of schooling across municipalities and years for this purpose. The CASEN survey is a nationally representative household survey which reports the tuition costs actually paid by families at high school and college levels and any amount of subsidies received, for each of the years 1998, 2000, 2003, 2006 and 2009. I use this data to retrieve the total tuition fees the family would have to pay at each education level, and I construct a tuition cost index by municipality and year at each level of education, which is incorporated as a proxy for the monetary costs in the model simulations. I take out time and municipality dummies from the construct in order to control for concurrent trends with wages and labor supply.

$$TC_{t,R}^{Ed} = \delta_0^{Ed} + \delta_1^{Ed} * R^{Ed} + \delta_2^{Ed} * t + \nu_{t,R}^{Ed}$$

so the average residual by time is then used to simulate schooling choices

$$U_{k,R}^{Ed} = \gamma_{1,k}^{Ed} \mu_k - \nu_R^{Ed} - \eta^{Ed}$$

Finally, in order to identify transition costs across sectors, I exploit the variation of mobility rates by schooling that I observe in the data.

### 2.5.3 Estimation

I estimate the model by Indirect Inference (Gourieroux, Monfort, and Renault (1993)). Meghir and Rivkin (2010) emphasize the use of simulation methods for structural estimation, as they do not use all of the information and restrictions implied by the model, given the available data, as MLE methods do, thus speeding



up the estimation process. The accuracy of the estimated parameters depends only on good specification of the data identifying moments, which is relatively simple in linear models.

The idea of indirect inference is to simulate data with the model at each iteration of the vector of structural parameters  $(\theta)$ <sup>10</sup>The process starts by simulating data from an initial vector of structural parameters  $(\theta_0)$ , and passing both the actual and simulated data by an auxiliary model, usually a system of linear regressions, to generate a set of data auxiliary parameters  $\beta$ , and the analogous set of simulated auxiliary parameters  $\beta(\theta)$ . At each iteration of the structural parameters  $\theta_j$ , Indirect Inference optimally finds the estimates of  $\theta_{j+1}$  that minimize the distance between the data and simulated auxiliary parameters, until convergence is achieved. For example, in the first iteration the set of initial simulated auxiliary estimates is  $\beta_1(\theta_0)$  and the following set of converging parameters  $\theta_1$  is found by minimizing a weighted distance between the simulated and data auxiliary parameters. The objective function at given iteration  $j$  is given by the metric

$$\text{Min}_{\theta}(\beta - \beta_j(\theta))'W(\beta - \beta_j(\theta))$$

where  $W$  is the optimal weighting matrix. Given the large number of moments required to identify the model, I use the diagonal of the optimal weighting function defined by  $\hat{W} = \text{diag}(VCV(\beta)^{-1})$ , which is obtained from the auxiliary estimates, and is sufficient to obtain consistent estimates. The use of a non-efficient weighting matrix has implications for the estimation of standard errors that are explained below.

A standard selection problem involves the estimation of a set of auxiliary linear regressions including log wage regressions and LPM models for participation both with the actual data and the simulations. Building on this approach, I use the following auxiliary model

<sup>10</sup>Simulating moments involves a forward recursive data generation process which uses as inputs the initial guess of the parameter vector  $(\theta_0)$ , a random draw of the vector of shocks for every individual at every age  $(\eta_{i,t}, \epsilon_{i,t})$ , and the optimal policy function  $d_t^*$ . The forward recursion process works as follows: in period  $t=1$  the optimal choice is retrieved by evaluating the policy function in  $\Omega_1$ , whereas the simulated counterfactual wages are obtained by evaluating  $W^m(\theta_0, \epsilon_1, \Omega_1)$ . The optimal choice  $d_1^*$  involves individuals choosing either education, work in sector  $m$  or unemployment in the first period, so the state space is updated accordingly for the next period accumulating either education or sector experience, and  $\Omega_2$  is evaluated for each simulated individual. This process is repeated until the whole sequence of choices  $d_t^* = [d_1^*, \dots, d_T^*]$  and counterfactual wages  $W_t^m = [W_1^m, \dots, W_T^m]$  are obtained and used to simulate moments. I simulate data on choices and counterfactual wages for  $N = 10,000$  individuals in  $T = 52$  periods.

$$\begin{bmatrix} \ln W_{it} \\ P(d_{it} = J) \\ P(d_{it} = J | d_{it-1} = J') \\ \Delta \ln W_{it}^{JJ'} \end{bmatrix} = Z'_{it} \delta + \nu_{it} \sim N(0, \Lambda)$$

which involves estimating a log wage regression, the probability of participation in sector  $J$ , the transition probability from sector  $J'$  to sector  $J$ , and the growth of log wages across sectors, on a vector of observable variables  $Z_{it}$  which includes the schooling level, sector experience, age, and tuition costs. Additionally, we must include in the set of moments the time series properties of the log wage regressions by sector, namely the autocorrelation and the VCV matrix of the shocks innovations. The vector of auxiliary parameters  $\beta$  to be matched are the coefficients of the auxiliary regressors ( $\delta$ ), and the VCV of the residuals  $\Lambda$ .

## Asymptotic Properties

Standard errors of the estimates are obtained by applying the asymptotic properties of GMM estimators described in Hayashi (2000). Let  $\hat{W}$  be a symmetric and positive definite weighting matrix such that  $\hat{W} \rightarrow W$  as  $N \rightarrow \infty$ , and let  $\theta(\hat{W})$  the set of estimated parameters. Under standard regularity conditions it can be shown that the asymptotic properties of the estimators are described by

$$\sqrt{NT}(\theta(\hat{W}) - \theta_o) \longrightarrow N(0, avar(\theta(\hat{W})))$$

A consistent estimate of the asymptotic variance is

$$\widehat{avar}(\theta(\hat{W})) = (d' \hat{W} d)^{-1} d' \hat{W} \hat{S} \hat{W} d (d' \hat{W} d)^{-1}$$

where  $\hat{S}$  is a consistent estimate of the VCV matrix of the data auxiliary parameters and where  $d = \frac{\partial \beta}{\partial \theta}$  is the gradient of the objective function evaluated at the vector of estimated parameters.

The matrix  $\hat{S}$  is obtained by bootstrapping methods and the gradients  $d$  are obtained by simulation. Each of the estimated structural parameters is shocked a sufficiently small value ( $\varepsilon$ ), and partial derivatives are obtained by

$$d = \frac{\beta(\theta(\hat{W}) + \varepsilon) - \beta(\theta(\hat{W}))}{\varepsilon}$$

## Smoothing the Objective Function

Using Indirect Inference in discrete choice models imposes important challenges. As Magnac, Robin, and Visser (1999) note, in discrete choice environments, objective functions are step functions of the structural parameters, which makes the use of derivative-based methods difficult. Derivative-based methods are generally preferred to local or global search methods because of speed and accuracy considerations. I use the approach presented by Keane and Smith (2003) who propose the use of a smoothing function allowing estimation by gradients.

To correct for the choppiness of the objective function, they propose an alternative system of auxiliary regressions to be used in the simulated data, which consists in proxying  $d_{it}^J$  by a smooth function of the structural parameters obtained from simulated value functions

$$g^J(\theta) = \frac{\exp(V^J(\theta)/\lambda)}{\sum_j \exp(V^j(\theta)/\lambda)}$$

where  $g^J(\theta)$  can be interpreted as the asymptotic probability of choosing alternative  $J$ , and  $\lambda$  is a calibrated smoothing parameter. The mirroring system of auxiliary regressions then becomes

$$\begin{bmatrix} \sum_J g_t^J * \ln W_{it}^J \\ g_t^J \\ g_t^{JJ'} \\ g_t^{JJ'} * \Delta \ln W_{it}^{JJ'} \end{bmatrix} = Z_{it}' \delta(\theta) + \nu_{it} \sim N(0, \Lambda)$$

Note that in the simulated auxiliary system we can observe each of the counterfactual wages. Therefore, the simulated log wages are the expected log-wages across sectors.

## Identifying Moments

In order to gain identifying power of the unobserved heterogeneity parameters, I add to the set of auxiliary parameters the proportions of people below wage percentiles {1.0, 25, 50, 75 and 90} by education level and sector, and regressions of log wages by sector on education, experience and age (Blundell, Costa-Dias, Meghir, and Shaw (2013)). The set of moments involves 155 auxiliary parameters used to estimate 45 structural parameters.

Structural Parameters	Identifying Moments
A. Wage returns	Log wage regressions by sector on $\{Ed, X^m, age, TC\}$
B. Shocks	Autocorrelation and variance of innovation of log wage residuals
C. Sector preferences	Participation regressions on $\{Ed, X^m, age, TC\}$
D. Transition Costs	Wage growth across sectors on $\{Ed, X^m, age\}$ Transition probabilities regressions on $\{Ed, X^m, age\}$
E. Type-specific parameters	Quantiles Log wages by sector and schooling

Table 2.3: Set of identifying moments

Table 2.3 describes the set of matching moments linked to the set of structural parameters I attempt to identify.

Finally, besides the discussion on the model identification, I check empirically whether the set of moments proposed actually identifies the set of structural parameters. I perform Montecarlo simulations using the model to generate an artificial dataset from an arbitrary set of structural parameters, which for this exercise are the “true” parameters. Once the artificial data is generated, I conduct the estimation procedure starting from an initial guess, 20%, 40% and 100% deviated from the “true” parameter and check for convergence. In Appendix B I show that the strategy for model identification combined with a choice of a set of identifying moments do a good job in identifying the “true” structural parameters.

## 2.6 Results

In what follows I present the estimation results organized in three sections. The first section shows the Goodness of Fit, validating the model performance in the replication of data patterns. In the second section I discuss the structural estimates. I use these results to answer the first two research questions I attempt to address, that is to say, the relative importance of preferences and wage returns for choices, and the extent to which skill functions are different across the formal and the informal sectors. In the third section, I use the model estimates to assess how individuals react to incentives in a dynamic context. By performing two simulation exercises, here I address the third research question of the paper evaluating the effects of labor market expectations on choices. Moreover, as the counterfactual simulations are

based on recently implemented policies by the Chilean Government, the predictions are informative about the potential effects of those policies on schooling attendance and the size of the formal sector.

### 2.6.1 Goodness of Fit

First I evaluate the model fit. Figures 2.5, 2.6 and 2.7 show the model fit for gross log wages. Overall, simulated wages do a good job in replicating wage profiles by schooling for the estimated sample, even though the fit is better for high school degree and college levels. The sample data is added for reference in dotted lines. In the case of less than high school level, there is a lack of data availability for informal workers younger than 18 years old, which may be the reason why informal wages for young workers are under predicted, while informal wages for older workers are over predicted. A similar situation occurs with informal wages for college level, even though in this case the model simulations seem to fit the general data pattern.

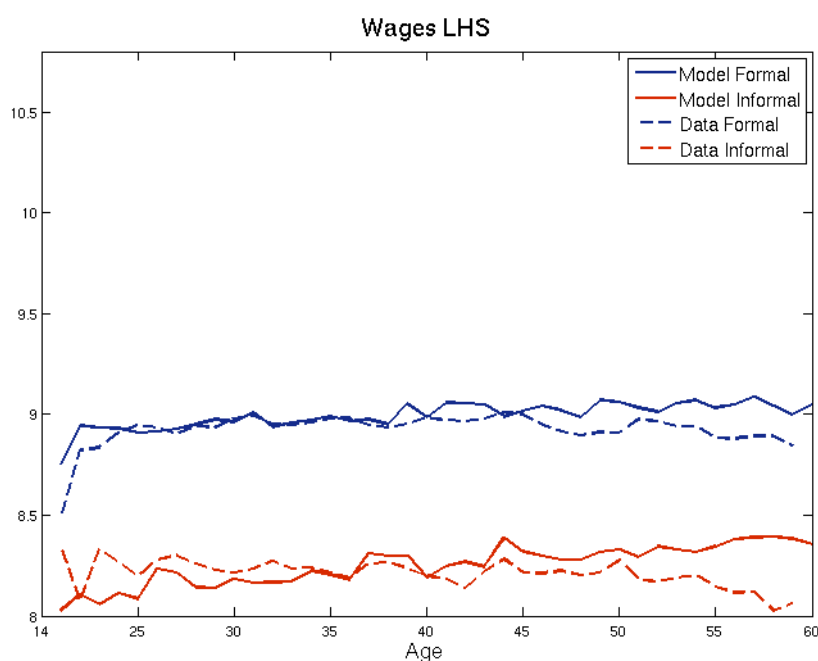


Figure 2.5: Model fit wages Less than High School level

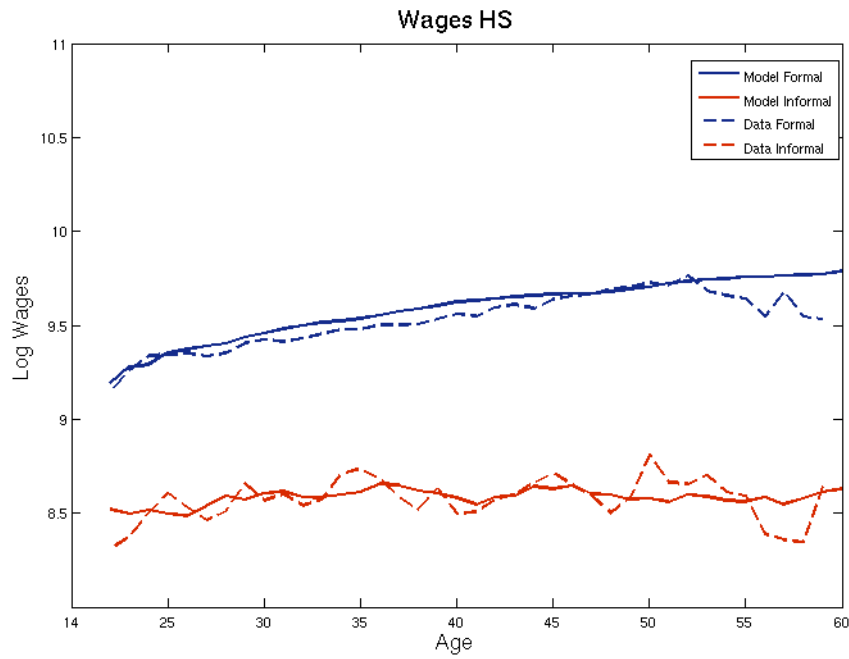


Figure 2.6: Model fit wages High School Degree level

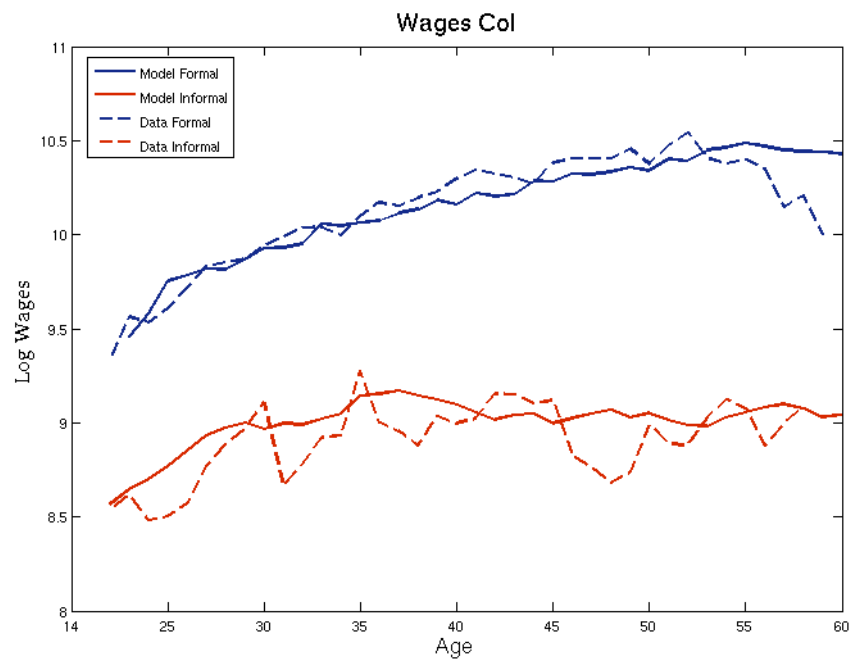


Figure 2.7: Model fit wages College Level

Figure 2.8 shows the data and simulated informality rates by schooling. Overall, the model seems to do a good job in fitting the general profile of sector participation

by schooling, preserving both the rank and life-cycle profiles of data patterns as shown in Figure 2.2.

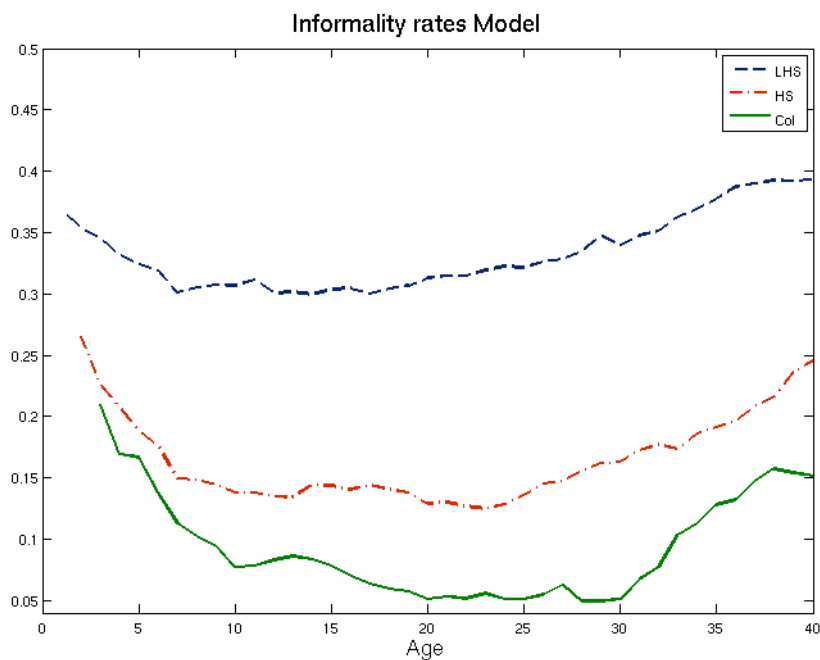


Figure 2.8: Simulated Informality Rates by education

Finally, Figures 2.9 and 2.10 report the model fit for unemployment/leisure/home production and schooling participation. Model simulations of school attendance show a good fit to the data. The model predicts that 37.9% of the simulated sample finished secondary schooling, while this is true for 36% in the sample. In the same way, the model predicts that 27.5% attended at least two years of college, while this is true for 28.8% of the sample. With regard to non-labor participation, the model is able to replicate the general patterns by age, but it somehow over-predicts the fraction of individuals who stay out of the labor market when they are older.



Figure 2.9: Model fit Unemployment/Home Production

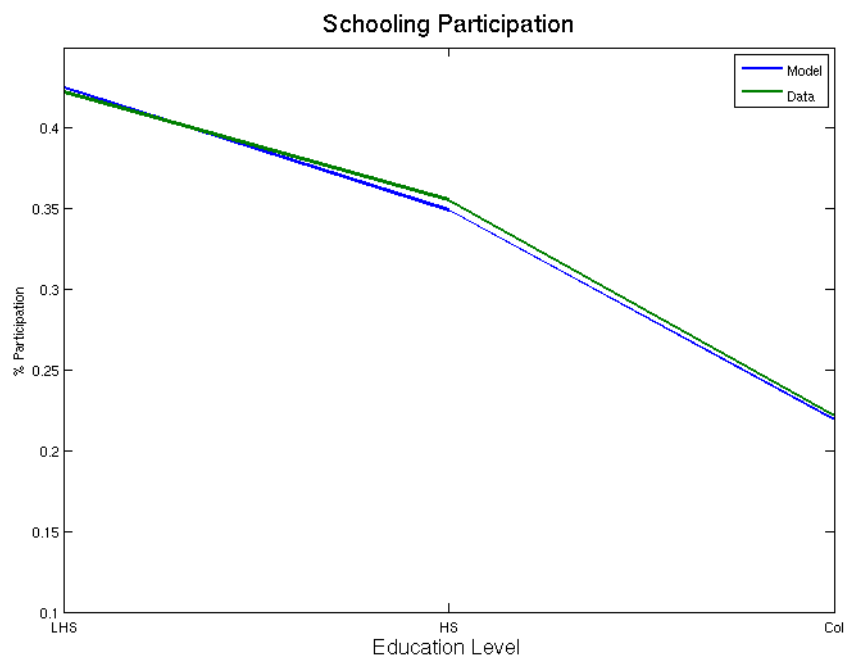


Figure 2.10: Model fit Schooling Participation



### 2.6.2 Structural Estimates

Table 2.4 shows the estimated preference parameters, which shed light about the relative importance of wage returns and preferences for sector amenities, and the estimated transition costs.

First, I assess the relative importance between wage returns and sector preferences by comparing wage valuations and fixed cost of working. Marginal valuation of incomes are statistically larger than fixed cost of work for individuals with Less than High School and High School degree, while they are similar for individuals with College level. Therefore, less educated people seem to be more liquidity-constrained, as their valuation for the cash-in-hand aspects of the job are more important. The trends go in opposite directions looking at fixed costs to work. Individuals at higher levels of education seem to value more the non-monetary aspects of the wage, no matter the sector in which they participate.

Second, the comparisons of these two parameters across sectors within education level brings important considerations. For Less than High School individuals, fixed costs to work are fairly similar across sectors, while they tend to value wage returns in informal activities more. These findings are consistent with previous evidence that low-educated informal workers tend to value the possibility of evading taxes and social security contributions from which they derive little value, and with evidence that this education group not only participates more of informal activities, but mobility rates between formal and informal activities are larger (Maloney (2004), Pagés and Stampini (2007)). Moreover, while marginal valuation of income is similar across sectors for individuals with High School degree and College level, fixed costs of work are consistently larger in the informal sector for these education groups. In summary, more educated workers face net costs in the informal sector, so the participation of these workers in informal activities should be explained by other reasons like returns to skills.

Third, predicted transition costs are larger for transitions towards the formal sector either from informal jobs or from non-participation. As these mobility costs have been estimated taking into account comparative advantage factors, preferences and shocks, I interpret these findings as evidence of the presence of some barriers to mobility to the formal sector.<sup>11</sup>

---

<sup>11</sup>Note that estimated switching costs are very large in all directions. Kennan (2008) argues that when there are few transitions in the data, estimated switching costs are implausibly large, because transitions must be attributed to unobserved payoff shocks. Therefore, he notes that observed switches must be attributed to unobserved payoff shocks and in order to evaluate their magnitude, one should evaluate how large the switching costs are conditional on the switch being made. When shocks are drawn from the type I extreme value distribution, he shows that the

Preference Parameters Working Sectors	Formal		Informal	
Wage Valuation ( $\gamma_{2,Ed}^m$ )				
LHS	1.31	(0.09)	1.41	(0.1)
HS	1.21	(0.1)	1.20	(0.09)
Col	0.83	(0.09)	0.86	(0.1)
Fixed Costs ( $\gamma_{3,Ed}^m$ )				
LHS	-0.62	(0.07)	-0.63	(0.15)
HS	-0.73	(0.04)	-0.82	(0.04)
Col	-0.74	(0.05)	-0.89	(0.05)
Transition Costs ( $c^{ij}$ )				
Formal to	-		-0.81	(0.09)
Informal to	-0.98	(0.04)	-	
Unemployment to	-1.06	(0.03)	-0.15	(0.08)
Linear age effects ( $\gamma_4^m$ )	0.22	(0.04)	0.15	(0.03)
Quadratic age effects ( $\gamma_5^m$ )	-0.0046	(0.0005)	-0.0039	(0.0005)

Table 2.4: Estimated Preference parameters Working Sectors

Table 2.5 shows the estimated type-specific parameters, which provides heterogeneity in comparative advantage within the model, and allow to relate schooling and labor choices. Consumption value of schooling is the underlying ability driving schooling attainment, while the returns to initial endowments represent how these abilities are rewarded across sectors. Among the unobserved types, type 2 is related to higher levels of ability and accounts for 84% of the sample, as they face a lower consumption value of schooling net cost of effort both at high school and college levels. Moreover, initial endowments have higher wage returns in the formal sector for both types. However, the wage premium in the formal sector is larger for type 2 than for type 1. As type 2 individuals are the high ability types, one can conclude that there is a positive association between ability and the formal wage premium to initial endowments. A natural implication of this finding is that workers with higher levels of ability, the ones who self-select into more schooling, are also the ones who self-select more into formal activities as these abilities are better rewarded in this sector.

---

average transition costs of moving from sector  $i$  to  $j$  are the estimated costs net of the difference in payoff shocks,

$$AVCost^{ij} = TR^{ij} - E[\eta_t^j - \eta_t^i \mid d_t^j = 1]$$

As a result, the monetary magnitude of the estimated transition costs is much smaller once they are adjusted by shocks.

Unobserved Heterogeneity	Structural Parameter	Type 1		Type 2	
Consumption Value Schooling ( $\gamma_{1,k}^m$ )	HS	-1.18	(0.09)	-0.35	(0.06)
	Col	-1.28	(0.1)	-0.42	(0.05)
Returns to Initial Endowments ( $\alpha_{0,k}^m$ )	Formal	8.32	(0.4)	8.97	(0.25)
	Informal	7.69	(0.2)	7.31	(0.4)
Probability Type 1		0.84 (0.002)			

Table 2.5: Estimated Preference parameters Non-working choices

Table 2.6 shows the estimated returns to schooling and experience across sectors. Returns to high school degree are larger in the formal sector, but I find a wage premium in the informal sector for college level. This result is in line with previous evidence for LAC countries (Amaral and Quintin (2006)). Arbex, Galvao, and Gomes (2010) argue that a wage premium in the informal sector for high-skilled workers has to be the reason why there is persistent participation in informal activities at college level, as these workers must somehow be compensated for giving up larger benefits in the formal sector. This argument is consistent with my finding that more educated workers value relatively more the non-wage attributes of formal jobs compared to informal jobs, but they participate in the informal sector as a consequence of larger returns to schooling. A second part of the explanation relies on composition effects. In my sample, more than 90% of informal workers over 40 years old and with post-secondary education are self-employed. Therefore, it is fair to conclude that the premium is largely attributed to successful entrepreneurial activities.

Second, the estimated returns to sector-experience suggest that formal experience is greatly valued in the formal sector at all education levels, but it is also valued in the informal sector for the less-skilled. In contrast, informal experience is valued in the informal sector only for the low-skilled, and it seems to be detrimental in the formal sector. This is consistent with multiple evidence from LAC countries showing that transitions between informal and formal jobs are larger for the less-skilled (Bosch, Goni, and Maloney (2007)). Finally, the estimated distribution of sector wages shows that productivity shocks are positively correlated across sectors, and that the variance of informal wages is larger, presumably because of less availability of wage data in this sector.

Wage Returns	Structural Parameter	Formal		Informal	
Returns to Schooling ( $\alpha_1^m, \alpha_2^m$ )	High School	0.38	(0.025)	0.35	(0.03)
	College	0.70	(0.03)	1.02	(0.03)
Returns to Formal Exp ( $\alpha_{3,Ed}^m$ )	LHS	0.14	(0.01)	0.06	(0.008)
	HS	0.13	(0.009)	-0.06	(0.009)
	Col	0.06	(0.01)	-0.05	(0.01)
Returns to Informal Exp ( $\alpha_{4,Ed}^m$ )	LHS	-0.06	(0.01)	0.10	(0.01)
	HS	-0.047	(0.01)	-0.02	(0.009)
	Col	-0.009	(0.008)	-0.05	(0.009)
Shocks	Autocorrelation ( $\rho^m$ )	0.91	(0.04)	0.87	(0.09)
	Std. innovation ( $\sigma^m$ )	0.25	(0.03)	0.27	(0.07)
	Correlation	0.32 (0.03)			

Table 2.6: Estimated Wage Offers

## 2.7 Studying Segmentation and Dynamics

The structural estimates are used to answer two empirical questions. First, they are used to assess the importance of labor market segmentation by comparing the relative weights of human capital, preferences and mobility costs in determining transitions from and towards the formal sector. Second, I simulate the effects of recently implemented schooling and labor subsidies to assess the empirical importance of labor market expectations and dynamics.

### 2.7.1 Are labor markets segmented?

Magnac (1991) defines labor market segmentation as a characteristic of dual labor markets in which the rewards in different economic sectors may differ for workers with equal potential productivity and the entry of workers to the formal sector is rationed. One way in which the model can be used to assess the degree of labor market segmentation is by assessing the relative importance of mobility costs with respect to human capital and preferences in determining changes in informality rates.

For illustration, suppose that labor markets are dual and there is no unemployment. In that case, the probability of switching to the formal sector for an individual of a given level of ability, education, accumulated experience and age can be represented by

$$P(d_t = F | d_{t-1} = I) = \Lambda \left\{ \left[ W_{t,k}^F (X_t^F + 1) - \gamma_{2,Ed}^m W_{t,k}^I (X_t^F + 1) \right] - \left[ \gamma_{3,Ed}^F - \gamma_{3,Ed}^I \right] - c^{I,F} + \beta \left[ Emax[V_{t+1}^F, V_{t+1}^I | d_t = F] - Emax[V_{t+1}^F, V_{t+1}^I | d_t = I] \right] \right\}$$

where  $\Lambda$  is the cumulative multinomial logit density function. In this expression, transitions are driven by four elements. The first bracket represents the contribution of higher accumulated experience in the formal sector that potentially pays off in both sectors. These gains are positive as the estimated returns to formal experience are larger in the formal sector (Table 2.6). Second, the contribution of larger job amenities in the formal sector at different levels of education. Estimates of  $\gamma_{3,Ed}^m$  in Table 2.4 show that fixed costs to work in the formal sector are smaller in the formal sector. Third, workers pay the estimated transition cost from the informal to the formal sector  $c^{I,F}$ . And finally, the dynamic effects resulting from the internalization of expected future payoffs of current decisions.

Table 2.7 presents the elasticities of transitions to the formal sector (or the reduction in informality rates) as a result of: (a) an 10% increase in the formal wage; (b) a 10% increase in formal job amenities (or a 10% increase in  $\gamma_{3,Ed}^F$ ), and (c) a 10% reduction in the transition costs towards the formal sector  $c^{I,F}$ . In average, changes in preferences for job amenities produce the largest reduction in informal labor participation, followed by mobility costs and wages. Individuals with Less than High School are the most responsive. In average, individuals with LHS level would decrease informality rates by 4.1% in response to a 10% reduction in mobility cost to the formal sector, a decrease of 12% in response to a 10% increase in formal sector amenities, and a reduction of 2.2% in response to a 10% increase in the formal wage. Individuals with HS degree would reduce participation in the informal sector by 2.5%, 9.1% and 1.8% respectively, and individuals with College education by 1.5%, 7.5% and 1.0%.

Age Group	10% Increase in Formal Wage			10% Increase in Job Amenities Formal Sector			10% Decrease in Mobility Costs towards the Formal Sector		
	LHS	HS	Col	LHS	HS	Col	LHS	HS	Col
14-17	0.1%	-	-	6.3%	-	-	1.7%	-	-
18-22	0.2%	1.5%	-	7.0%	11.5%	-	1.8%	1.9%	-
23-26	3.1%	1.6%	0.5%	9.4%	10.6%	7.5%	3.1%	2.1%	1.0%
27-30	3.7%	1.8%	1.1%	10.1%	8.8%	7.7%	3.5%	2.3%	1.1%
31-35	3.2%	1.8%	1.9%	12.2%	7.9%	6.5%	3.6%	2.6%	1.4%
36-40	2.9%	1.9%	1.3%	13.7%	8.1%	7.1%	3.9%	2.6%	1.5%
41-45	2.3%	1.9%	0.7%	15.1%	8.4%	7.6%	4.9%	2.8%	1.6%
46-50	1.8%	1.8%	0.6%	16.1%	8.8%	7.8%	5.9%	2.9%	1.6%
51-55	2.1%	2.0%	1.1%	15.9%	9.4%	7.8%	6.2%	2.8%	1.7%
>55	2.9%	2.1%	1.1%	13.9%	8.6%	8.0%	6.2%	2.9%	1.7%
<b>Total</b>	<b>2.2%</b>	<b>1.8%</b>	<b>1.0%</b>	<b>12.0%</b>	<b>9.1%</b>	<b>7.5%</b>	<b>4.1%</b>	<b>2.5%</b>	<b>1.5%</b>

Table 2.7: % Change in informality rates due to changes in wages, preferences and mobility costs

I use these simulations to define labor market segmentation as the relative effect of mobility costs with respect to comparative advantage (wages + preferences) in explaining informality rates. Table 2.8 shows that mobility costs explain 28.1%, 23.6% and 16.9% in the variation of informality rates respectively for LHS, HS and individuals with College education. In conclusion, barriers to mobility unexplained by human capital accumulation and preferences are not as important as comparative advantage in driving labor market participation, and they are decreasing in education.

Age group	Labor Market Segmentation		
	LHS	HS	Col
14-17	26.5%	-	-
18-22	25.0%	14.7%	-
23-26	24.8%	17.3%	12.5%
27-30	25.3%	21.5%	12.6%
31-35	23.4%	26.7%	16.8%
36-40	23.2%	26.1%	17.8%
41-45	28.5%	27.2%	19.3%
46-50	33.1%	27.4%	19.0%
51-55	34.6%	24.5%	19.0%
>55	37.1%	27.2%	18.6%
<b>Total</b>	<b>28.1%</b>	<b>23.6%</b>	<b>16.9%</b>

Table 2.8: % weight of Labor Market Segmentation explaining transitions to the formal sector

## 2.7.2 The role of dynamics and labor market expectations

I use the model estimates in combination with recently implemented educational and labor policies to assess the importance of labor market expectations and persistency of choices in labor market participation. I first evaluate the incentives provided

by a revenue neutral 20% wage subsidy aiming at supporting the incorporation of disadvantaged workers between 19 to 26 years old in formal jobs.

Table 2.9 describes the effects of such a subsidy on schooling participation. Overall, an exogenous increase in monetary incentives to participate in the formal sector increases secondary schooling completion rates by 1.0%, but it slightly decreases the incentives for college participation by 0.5%. In the case of high school graduates, the interpretation is straightforward. Returns to secondary schooling are larger in the formal sector, thus the subsidy increases the incentives for formal labor participation as a high school graduate. In the case of college attendance, one plausible explanation for this result is that the subsidy provides incentives to dropout of schooling straight after high school completion and start working in formal activities. This is particularly true for type 2, the high ability type, who are likely to succeed anyway in the labor market even without a college degree.

Table 2.10 describes the effects on sector participation. Overall, informality rates decline substantially (2%) for workers within the targeted age group, and due to dynamics, the decrease in informality rates continue until the end of the life-cycle, although at lower levels. Informality decreases the most for the less-skilled (2.7%), who are likely to value the monetary aspects of labor market incentives more. Moreover, type 1 is more responsive to the tax reduction because the ability premium in the formal sector for this group, although positive, is lower than for type 2. Table 9 and Table 10 (Appendix C) simulate the effects of extending the tax reduction to the age of 40. In this scenario, the decrease in informality rates is larger and more permanent for both types.

The simulation exercise shows that individuals would contemporaneously react to future sector-specific labor market shocks and that their current choices would remain persistent. This confirms the importance of labor market expectations and dynamics in participation in informal labor markets.

Schooling Participation	HS	Col
Total Sample	1.0%	-0.5%
Type 1 (low)	0.8%	-0.3%
Type 2 (high)	1.9%	-2.8%

Table 2.9: Effects of a 20% wage subsidy to formal employment on schooling

Age	Total Sample			Type 1			Type 2		
	LHS	HS	Col	LHS	HS	Col	LHS	HS	Col
14-17	-2.3%	-	-	-2.1%	-	-	-4.4%	-	-
18-22	-3.0%	-2.1%	-	-2.9%	-1.7%	-	-4.4%	-1.8%	-
23-26	-2.3%	-1.7%	-1.2%	-2.3%	-1.4%	-1.7%	-0.6%	-0.7%	-1.1%
27-30	-1.7%	-1.2%	-0.4%	-1.8%	-0.8%	-0.7%	-1.4%	-0.5%	0.0%
31-35	-1.1%	-0.5%	-0.4%	-1.2%	-0.3%	-0.7%	-0.2%	-0.4%	0.1%
36-40	-0.9%	-0.2%	-0.7%	-1.1%	-0.2%	-1.0%	-0.3%	-0.1%	0.1%
41-45	-0.8%	-0.1%	-0.6%	-0.9%	-0.1%	-0.8%	-0.5%	-0.1%	0.1%
46-50	-0.4%	0.0%	-0.7%	-0.6%	0.0%	-1.0%	-0.6%	0.0%	0.1%
51-55	-0.3%	0.0%	-0.4%	-0.5%	0.1%	-0.6%	-0.5%	-0.1%	0.1%
>55	-0.3%	-0.1%	-0.2%	-0.5%	0.0%	-0.3%	-0.4%	-0.1%	0.2%

Table 2.10: Effects of a 20% wage subsidy to formal employment on informality rates

Finally, Table 2.11 assesses the redistributive effects of the subsidy. If the subsidy was allocated to all workers belonging to the targeted ages and the cost was shared on a per-capita basis, the subsidy would have detrimental effects, increasing earning inequalities. Type 2 would benefit the most from the increase in monetary incentives because this type has the largest ability premium in the formal sector, so they would be formal anyway. In contrast, if the government was able to observe types and could target the subsidy only at Type 1, there would be a positive redistributive effect of such a policy.

In reality, types are not observable. One alternative to be explored is to exploit further the data on the workers' socio-economic background when they were children. Individuals are required to report whether the family was poor or not poor when they were young, the education of their mother and the education of their father. If one estimates the probability of being certain type conditional on family background, it would be possible to evaluate the redistributive effects of the true means-tested subsidy.



Expected lifetime earnings at age 14	Overall	Type 1	Type 2
Baseline	186,843	148,366	388,847
Standard wage subsidy	198,063	157,387	411,615
Gross Gain	11,220	9,020	22,768
Net gain		-404	20,973
Standard wage subsidy only Type 1	194,420	157,387	388,847
Gross Gain	7,577	9,020	0
Net gain		2,656	-1,212

Table 2.11: Effects of a 20% wage subsidy to formal employment on earnings inequality

I analyze how people would react to a monetary subsidy of 40% to finance college education. Table 2.12 shows that such a subsidy would have a small positive impact on college participation, increasing attendance by 1.8%, where the high ability types would concentrate the largest increases (3.3%). This is strongly driven by the fact that initial endowments by unobserved types are the main factor driving selection into schooling.

College attendance	%
Total Sample	1.8%
Type 1 (low)	1.5%
Type 2 (high)	3.3%

Table 2.12: Effects of a 40% college subsidy on schooling

Table 2.13 shows the effects of the subsidy on informality rates for workers attending college. Interestingly, the monetary incentives do not have any effect on informality rates on average, but there are differences by ability type. Type 2, the high-ability type, increase their informal labor participation between 0.5% and 0.7% below age 30, in line with the findings that college returns in the informal sector are larger. However, Type 1 workers decrease their participation in informal labor markets by between 0.4% and 0.8% for the same age group. Limited impacts of college subsidies on career progression in different sectors have been also found in the US by Keane and Wolpin (1997). In their findings, they note that initial endowments at age 16 play a key role in determining selection into college and selection into different types of jobs, reducing the impact of exogenous monetary incentives

that are allocated long after the most critical periods of skill formation have already finished.

Age	Average	Type 1	Type 2
23-26	-0.1%	-0.4%	0.5%
27-30	-0.1%	-0.8%	0.7%
31-35	-0.1%	-0.5%	0.4%
>35	-0.1%	-0.3%	0.4%

Table 2.13: Effects of a 40% college subsidy on informality rates

## 2.8 Concluding Remarks

I contribute to the literature on labor informality in developing countries by studying whether comparative advantage determines participation in informal labor markets. I develop a life-cycle model building on a Roy Model extended to endogenous schooling decisions and also extended to compensating wage differentials, to analyze how individuals with heterogeneous skills and preferences choose schooling and sector participation based on their comparative advantage. The model is estimated by exploiting rich cross-sectional and time-varied Chilean panel data on wages, sector participation and school attendance.

My contribution to the literature also relies on the exploration of the different mechanisms by which comparative advantage might drive choices. First, a structural estimation allows me to disentangle the relative importance of wage returns and preferences for non-wage attributes in driving decisions. More traditional approaches used to test whether comparative advantage drives informal labor participation might have some limitations. For example, the study of wage differentials between the formal and the informal sectors does not account for the fact that individuals might self-select into jobs based on utility differences, which are not captured by wages. In addition, studies based on the analysis of mobility rates may not be able to separately identify whether choices are driven by comparative advantage or if they are a consequence of idiosyncratic or industry-specific shocks.

Second, some authors have emphasized that a key difference between the formal and informal sectors is that they are likely to have different production functions. While formal firms are associated with larger size, higher capital-labor ratios, and higher technology, and demand more skilled labor, informal firms tend to concentrate unskilled labor-intensive economic activities and operate at lower levels of productivity. I test for this assumption in a partial equilibrium approach, by estimating skill functions that may vary across sectors. I show that returns to abilities, schooling and sector-specific experience largely diverge across the formal and informal sectors. Third, I assess the extent to which labor market expectations matter for decision-making. Model simulations show that individuals react significantly to labor market expectations by making different decisions regarding their schooling and participation in informal labor markets when the relative monetary incentives between the formal and informal sectors change exogenously. The effects of these incentives are persistent over the life-cycle, which indicates the strong effects of dynamics through human capital accumulation. Finally, I test for the existence of barriers to mobility by estimating transition costs that cannot be explained by skills, tastes or shocks.

The modeling framework I propose incorporates some innovations. To my knowledge, there is no previous literature that has studied participation in informal labor markets using a life-cycle approach, in which the schooling decision is endogenous. Moreover, the previous literature studying labor informality has largely stressed the importance of heterogeneity in skills and preferences shaping comparative advantage. I explicitly model unobserved heterogeneity by allowing initial endowments to jointly drive selection into schooling and sector productivity.

Some estimation results are important to highlight. First, comparative advantage is more important than labor market segmentation in driving participation in informal labor markets. Model simulations show that mobility costs would explain up to 25% of transitions to the formal sector while the rest is explained by human capital accumulation and preferences for job amenities. The results also show that there is a large heterogeneity in education. Less educated individuals attach more value to monetary rewards than non-wage attributes, and the relative valuation of wages for this group is larger in informal activities. These findings are consistent with previous evidence that less educated informal workers tend to value the possibility of evading paying taxes and social security contributions from which they derive little value. Instead, more educated individuals tend to attach more value to the non-wage attributes of jobs relative to wage returns, and these benefits are larger in the formal sector.

Second, the estimated sector-specific technologies of skill accumulation support the hypothesis that the formal and informal sectors have different production functions. I find that workers with higher levels of ability are better rewarded in the formal sector; returns to High School are larger in the formal sector; and, I find a wage premium in the informal sector for post-secondary schooling. The latter result explains why highly skilled individuals are found to participate in informal activities despite the fact that, by so doing, they give up the large non-wage benefits attached to formal contracts. Furthermore, in my sample, more than 90% of the highly skilled informal workers over the age of 40 are self-employed, which suggests that highly educated informal workers are actually people involved in successful entrepreneurial activities.

Third, I use the model estimates to assess the importance of labor market expectations for schooling and labor decisions. I show that a revenue neutral wage subsidy to formal youth employment would decrease informality by 2% for the targeted age groups, but it would also persistently decrease informality rates for older workers not affected by the policy, which is a consequence of the dynamic effects

of the incentives through sector-specific human capital accumulation. Furthermore, the subsidy would also increase the incentives to finish High School (1.5%), and slightly decrease the incentives for college attendance. I also find that the incentives provided by a college subsidy of 40% would increase college participation by 1.8%, with high-ability types being the ones who react the most to the incentives (3.3%). On average informality rates would not be affected, but the policy would significantly reduce informality rates for low-ability types.

And fourth, the estimated transition costs from the informal to the formal sector are substantially larger than in the opposite direction, while the re-entry costs from non-labor participation in the formal sector are also larger. Given that, in general, workers prefer to work in formal activities but transitions to this sector are more costly, I interpret these findings as evidence of the existence of some barriers to entry to the formal sector.

Finally, my research agenda incorporates two model changes intended to improve the understanding of the effect of comparative advantage on labor informality. The first improvement relates to the incorporation of self-employment as a third sector. As explained above, informality rates are larger for the elderly because of a composition effect. While informality for salaried employees decreases over the life-cycle, among the elderly a larger proportion are self-employed. Figure A.4 in the Appendix shows that if all of the self-employed are now considered as another sector, the economy is composed of the salaried formal, the salaried informal and the self-employed, and a clear rank emerges in terms of wage profiles. Remarkably, the formal salaried face higher wage returns than those in the other two sectors, with slopes increasing in schooling. And while, for individuals with less than High School level education, the wage profiles for the informal salaried and self-employed are similar, at higher schooling levels the self-employed become more similar to the formal salaried. This suggests that the informal salaried and the self-employed have different skill accumulation processes, so it may be worth modeling separately the three different sectors. The estimation of such a model is currently a work in progress.

The second improvement relates to the relaxation of some model assumptions. As discussed above, risk neutrality is a relatively safe assumption if one is willing to analyze the dynamics for young workers. However, an interesting extension of the model is the incorporation of risk aversion alongside both private and pension savings to analyze how schooling and labor supply will change in response to credit market imperfections. Additionally, in the current modeling framework, I have also assumed that individuals foresee perfectly, the returns to human capital accumulation across

sectors. While most of the literature on dynamic models builds on this assumption, recent modeling frameworks like the ones proposed by Altonji, Blom, and Meghir (2012), emphasize the importance of incorporating the more realistic assumption that individuals face uncertainty about their own preferences and the returns to human capital accumulation. I intend to incorporate richer sources of dynamics in wages and in the paths of human capital accumulation by modeling individuals that learn their own preferences and the technology of skill formation by doing.

# Chapter 3

## Does the Timing of Parental Income Matter?

### 3.1 Introduction

There is a large empirical literature examining the intergenerational transmission of economic status (for recent surveys see Solon (1999), Black and Devereux (2011), Björklund and Salvanes (2010)). It is possible to find estimates of intergenerational mobility for various outcomes for virtually every country in the world where data linking parents and children is available. Most estimates come from simple models linking a measure of child income and a measure of parental income, where incomes at a particular parental age window or life-time earnings are used.

$$Y_i = \alpha + \beta I_i + u, \quad (3.1)$$

where  $Y_i$  is a measure of the child's income,  $I_i$  is a measure of parental income, and  $u$  is a residual.

Standard theoretical models of intergenerational transmission justify the use of equation (1) (e.g., Becker and Tomes (1979), Becker and Tomes (1986)), but they usually collapse the childhood years to a single period of life. More realistic models of parental investments in children distinguish several stages of childhood (Cunha and Heckman (2007), Cunha, Heckman, and Schennach (2010)). They point out that the whole history (in particular, the timing) of parental investments in children may be as or more important than the total amount invested during the childhood years. Therefore, if there is a link between shocks to family income and investments in children, a simple model of parental income at one point in time may be misspecified.

This paper extends the literature on intergenerational transmission by examining

the relationship between adult outcomes of children and the timing of parental income during their childhood years. We use data from Norway for children born during the 1970s, which allows us to link an individual's outcomes as a young adult to the whole history of parental income during the childhood and adolescence years.

We address the three research questions. First, we test for the empirical importance of the timing of parental income relative to permanent income for human capital formation of their children. In doing so, we rely on rich administrative data, which allows us to estimate very flexible specifications without needing to assume strong parametric assumptions. Second, although causality of the timing of income on human capital is hard to be argued for several reasons, we attempt to address each of those concerns presenting a series of robustness checks. An third, we attempt to understand the mechanisms by which the timing of income shocks matter for human capital formation of children, by developing and estimating simple models of parental investments in children where parents face different sources of uncertainty and they are not perfectly insured against shocks.

Our sample consists of all individuals born in Norway between 1971 and 1980. It is possible to link each individual to his mother and father and their respective annual income for all years in this decade. We use this information to construct maternal and paternal income histories from the birth of the child until the year she was 17,<sup>1</sup> which are linked to a set of human capital outcomes available for their children in their 20's (years of schooling, high school dropouts, college participation), We also present the results for other available outcomes like IQ, income early in their careers, a health index, and fertility.

Ideally, we would want to estimate flexible functions of outcomes in adulthood on the series of annual incomes between the ages of 0 and 17 of the child. In practice it is difficult to implement such an estimator when the number of regressors is as high as this (18). Therefore we group the childhood years into three periods: ages 0-5, 6-11, and 12-17. We construct the average deflated and discounted income for each of these age groups, as well as a measure of permanent income during childhood which takes the sum of discounted income over the three childhood periods. We then estimate non-parametric regressions of each outcome on permanent income and incomes in two out of the three periods of childhood (since the third would be collinear), and semi-parametric models which allow the inclusion of parental characteristics as controls.

We present our results through a series of two dimensional graphs. Take for

---

<sup>1</sup>In this paper, the 0-17 age interval constitutes childhood.



example the case where we include as regressors permanent income, income at ages 6-11, and income at ages 12-17. We can fix, for example, permanent income and income at ages 12-17, and see how changes in income at ages 6-17 translate into changes in outcomes. This corresponds to the impact of shifting income between ages 0-5 and ages 6-11 on the adult outcomes of children.

The main finding of the paper is that, for a given level of permanent parental income, balanced income profiles lead to higher levels of education of the child than income profiles subject to several fluctuations. This is because it is better to smooth investments in children than to suffer large fluctuations (which are a consequence of income shocks). This result also holds for related outcomes like high school dropout and college enrollment. For other outcomes, such as earnings, IQ, or teenage pregnancy, the picture can be slightly different, indicating differences in the production function for different outcomes. In terms of magnitudes, although permanent income is still more relevant in the production function of human capital outcomes, the timing of income also matters. For example, increases in permanent income of £100,000 would increase child earnings by age 30 in 10% and would increase schooling attendance in 0.5 years, while if parents were able to shift the same amount from when the child is in middle childhood (6-11 years old) to early childhood (0-5 years old), earnings by age 30 would increase 5% and schooling attendance would increase by 0.25 years.

Our second finding is regarding causality. We realize that households jointly decide investments, consumption and labor supply, but investments are unobserved in this sample. Therefore, the timing of income can be a function of changes to labor supply for example because maternity leave choices for child nurturing, or because there are heterogeneous age-earnings profiles which depend upon parental human capital. We test for each of these concerns and find that the effect of the timing of income particularly on schooling is invariant to these additional controls.

In a third finding, we simulate models to explain why the timing of income matters. If parents face income uncertainty and borrowing constraints, they might not be able to smooth out income shocks, and they are transmitted into optimal consumption and investment decisions. If it turns out that early and late investments in human capital are complement in the technology of skill formation, then the inability to smooth income shocks would lead to underinvestments (see Cameron, Heckman, Journal, and April (1998), Cameron and Heckman (2001), Carneiro and Heckman (2002), Carneiro and Heckman (2003), or Cunha and Heckman (2007), among many others). We estimate a simple model of parental investment in children

emphasizing borrowing constraints and income uncertainty, and we show that the only technology of skill formation that is consistent with the data patterns is the one that is complementary in parental investments across childhood stages.

A few other authors have explicitly examined the role of the timing of income in the formation of human capital. Some of these focus on survey data from the US and Germany, and rely on relatively small datasets (Duncan, Yeung, Brooks-gunn, and Smith (1998), Levy and Duncan (2000), Jenkins and Schluter (2002), Carneiro and Heckman (2003), Caucutt and Lochner (2012)). Others use much larger register data for Denmark and Norway (Aakvik, Salvanes, and Vaage (2005), Humlum (2011)), but nevertheless they estimate very restrictive models. In particular, all these papers estimate regressions of child outcomes on the income of parents at different ages. Since the levels of income in different periods enter in a linear and additive way in these models, they are assumed to be “substitutes” in the production of human capital. Relative to the papers using US and German survey data our paper relies on much better data (larger samples and richer income histories), which allows us to estimate much more flexible models with considerable precision. This is also true when we contrast our analysis to the ones using register data for Norway and Denmark. The flexible models we estimate allow us to construct a much richer picture of the role of the timing of income than the one presented in previous work. This is important because the complementarity (or other interactions) of investments in human capital across periods (Cunha, Heckman, and Schennach (2010)) may translate into complementarity (or other interactions) of income shocks across periods.

The empirical importance of permanent income relative to the timing of income is likely to be dependent on the country that we study. In a country such as Norway in the 1970s, where the welfare state was starting to develop but already had well developed safety nets, it may be much easier to smooth fluctuations in income than in other countries where insurance possibilities are more limited, because the welfare system is not as generous, or capital markets are not as developed. The fact that income fluctuations matter at all in Norway, suggests that they may be even more important in poorer countries.

We realize that income fluctuations over the life-cycle are not entirely driven by shocks, but they also reflect parental choices. However, we provide robust evidence suggesting that the timing of parental income has long-term impacts on human capital formation. For example, family income during the first years of life of the child may be low because the mother decided to take maternity leave. Therefore, in

our main specification we use father's income alone, as opposed to mother's income or total family income. In addition we show that removing ages 0-2 from the data (and redefining the age groups to be 3-7, 8-12, 13-17) leads to similar results as in the main specification. The same happens when we include in the model controls for the proportion of years the mother is not working in each period.

We also realize that income profiles depend on parental human capital, which has an independent impact on child outcomes. Therefore we control for paternal and maternal education, which is allowed to interact with paternal and maternal age at birth. We also include controls for the slope of the income profile calculated using prebirth income and income occurring right after the 17th birthday of the child (outside the periods we use to construct income histories), so we can control for whether the father is on a high or a low income growth trajectory (or even on negative trajectory). Other controls include the birth year and gender of the child.

The combination structural estimation helps us to provide an interpretation of our results linking the empirical analysis with the economic theory. If markets were complete and there are no sources of uncertainty, then we should not find differential effects for the timing of income on human capital formation, even in the presence of strong complementarities of investments across different stages of childhood in the technology of skill formation. Parents could borrow and save as much as they want to smooth consumption and investments. However, if parents face borrowing constraints and they face a negative income shock they cannot borrow from future earnings or from child's earnings, so investments would not be optimal. Moreover, investments are reactive to income shocks in the presence of uncertainty, and savings alone do not provide perfect insurance. For example, if there is income uncertainty and investments are complementary, families with stable incomes do better as are able to keep stable investments.

We develop and estimate a model of parental investments in children with multiple periods of childhood, as in Cunha and Heckman (2007). In each period, parents decide consumption, savings and investments in children. Human capital is a general function of investments in different periods, so the whole history of investments potentially determines human capital formation (except in the special case where there is perfect substitutability between investments in different periods). Parents are subject to shocks to income, which can be permanent or transitory. Markets are incomplete, so there may be only partial insurance against shocks (as in, for example, Blundell, Pistaferri, and Preston (2008)). In sum, we add parental investments in children to a standard life-cycle model of consumption and savings,

with imperfect insurance. Finally, we also explore the role of other type of shocks, like uncertainty about the technology of skill formation, as it has been documented that parents who face uncertainty about their child ability and/or the returns to investing in their children, tend to delay those investments.

The structure of the paper is as follows. In section 3.2 we describe the data, and in section 3.3 we present our empirical methods. Section 3.4 discusses our results, and in section 3.6 we present simulations from dynamic models of parental investments in children which help us interpret the results. Finally, section 3.7 concludes.

## 3.2 Data

The data source is the Norwegian Registry data maintained by Statistics Norway for the periods 1971 up to 2006. It is a linked administrative dataset that covers the population of Norwegians and is a collection of different administrative registers providing information about month and year of birth, educational attainment, labour market status, earnings, and a set of demographic variables (age, gender) as well as information on families including parents' marital status. It is possible to link individuals to their parents, and to gather labour market information for both.

For the bulk of the analysis we select all births in the period 1971-1980. In particular, we construct annual paternal taxable earnings data for each year from the three years preceding the child's birth, through to their 20th birthday. An additional child outcome is available for births up to 1986, test scores in high school. Therefore, we map income data and parental characteristics for these additional cohorts of children. This gives us information on 522,490 children.

The earnings values include wages and income from business activity but also unemployment, and sickness benefits. Therefore, the selected income measures include some degree of insurance against low income shocks, i.e. when workers temporary out of work but still in the labor force, and consequently we expect the effect of the timing of taxable earnings to be lower than the effect of labour earnings alone (which we cannot measure). We discount all incomes to the year of birth of the child, using a fixed real interest rate of 4.26% (Aakvik, Salvanes, and Vaage (2005)). However, our results are robust to a wide range of real interest rates, and to time-varying discount rates.

In order to construct a measure of income in each of the three periods we take the average of discounted annual paternal incomes within each period (0-5, 6-11, 12-17). Permanent income is then defined as the sum of income in the three periods.

Even though the main outcome of interest of this research is years of education, we consider a large range of other child outcomes. The administrative data also includes schooling outcomes like an indicator for dropping out of high school at the age of 16, and college enrollment. Military service is compulsory in Norway for males, and between the age of 18-20 males usually take an IQ test. This test is a composite of arithmetic, words,<sup>2</sup> and a figures tests<sup>3</sup>, all of which are recognized as tests of IQ. We include also an indicator for teen pregnancy. This takes the value of 1 if the individual has a child aged between 16 and 20. We measure additionally

---

<sup>2</sup>The word tests are most similar to the Wechsler Adult Intelligent Scale (WAIS).

<sup>3</sup>The figures tests are similar to the Raven Progressive matrix

a health score taken from the military tests upon entry to the Army. This test is designed to ascertain physical capabilities of the males. It is measured on a 9 point scale, with the top score of 9 indicating health sufficient to allow military service. Around 85% of individuals have the top score.

Finally, we construct a set of control variables, which are fundamental in the discussion of the causality effects of the timing of income on human capital formation. First, we build a measure of the heterogeneous income profile of households, as the difference between income for child aged 18-20 and in the three years pre-birth<sup>4</sup>. This allows us to control directly for the slope of the income profile throughout the periods of childhood. Other controls include family background information of parental years of education and age at birth, marital status and family size in each year of the child's life. We observe also the year of birth of the child and the municipality of residence in each year of the child's life.

### 3.3 Methods

#### 3.3.1 Empirical Strategy

Let  $Y_i$  be an outcome of child  $i$  (education, high school drop out, college attendance, earnings, IQ, health, teenage pregnancy, grades in high school) in later adolescence or young adulthood. We are interested in  $Y_i$  as a function of the history of paternal income  $I_{it}$  in each period  $t$  ( $t = 1, 2, 3$ ), and permanent income of the parents,  $PI_i$ . Since  $PI_i = I_{i1} + I_{i2} + I_{i3}$  we drop one of the periods from the model, say  $I_{i1}$ . Therefore, we write:

$$Y_i = m(PI_i, I_{i2}, I_{i3}) + \varepsilon_i \quad (3.2)$$

We allow  $m(PI_i, I_{i2}, I_{i3})$  to be a non-parametric function of its arguments. The relationship between the timing of income and child outcomes needs to be flexible. Parents are faced with income shocks in each period and, in response, decide how much to invest in children (and how much to consume and save). There is a technology that links the adult human capital of an individual to the whole history of parental investments in childhood and adolescence. The link between income shocks and child outcomes, described by equation (3.2), depends on many factors, including preferences, technology, information, and the structure of credit markets (insurance possibilities). Therefore, this relationship can be quite complex.

---

<sup>4</sup>A robustness check conditions instead for the growth rather than the level of pre- and post-childhood income.

We are particularly interested in  $m_2(PI_i, I_{i2}, I_{i3}) = \frac{\partial m(PI_i, I_{i2}, I_{i3})}{\partial I_{i2}}$  and  $m_3(PI_i, I_{i2}, I_{i3}) = \frac{\partial m(PI_i, I_{i2}, I_{i3})}{\partial I_{i3}}$ .  $m_2(PI_i, I_{i2}, I_{i3})$  tells us the impact on outcome  $Y_i$  of shifting income from period 1 to period 2, since we are keeping  $PI_i$  and  $I_{i3}$  fixed (and  $PI_i = I_{i1} + I_{i2} + I_{i3}$ ). In our empirical section we will present a series of graphs relating  $Y$  and  $I_2$  (for different outcomes  $Y$ ). The graphs will vary depending on the values of  $PI_i$  and  $I_{i3}$  on which we evaluate this function. An analogous interpretation and graphical representations of results can be given to  $m_3(PI_i, I_{i2}, I_{i3})$ .

$\varepsilon_i$  should be interpreted as the unobserved heterogeneity that is left after controlling for permanent income in the model. Therefore, we assume that  $\varepsilon_i$  has a finite conditional variance:  $E(\varepsilon_i^2 | PI_i, I_{i2}, I_{i3}) \leq C < \infty$  and that  $E(\varepsilon_i | PI_i, I_{i2}, I_{i3}) = 0$ . We are interested not in the impact of  $PI$  itself on  $Y$ , but on the impact of the timing of income ( $I_2$  and  $I_3$ ) on  $Y$ , after conditioning on  $PI$ . In other words, we want to compare (the late adolescence or adult) outcomes of children whose parents have the same level of permanent income between the ages of 0 and 17, but differ in the level of income they get in each period.

We would like to interpret  $I_2$  and  $I_3$  as income shocks orthogonal to other determinants of outcomes  $Y$ , conditional on  $PI$ . It is likely that  $PI$  absorbs much of the relevant unobserved heterogeneity across parents (correlated with the overall level of their income), but one may still be concerned that parents facing different income profiles may be also different in many other dimensions.

In order to address this issue we start by excluding maternal income from the model, and rely only on paternal income to construct  $(PI_i, I_{i2}, I_{i3})$ . Maternal income in each period could be very much related to decisions of staying at home caring for children instead of work, which is likely to affect child outcomes (e.g., maternity leave; see Carneiro, Løken, and Salvanes (2011)). On the other end, paternal income is much less likely to be affected by these choices.

In addition, we condition on paternal education interacted with paternal age at birth (by including dummies for years of education and age at birth interacted with each other). This controls for different age-education profiles across fathers. Moreover, we construct a measure of paternal income growth between the ages of 0 and 17 of the child, based on income 1 to 3 years before birth (pre-birth income), and income 1 to 3 years after age 17 (post-17 income). This means that we explore fluctuations in income around deterministic age-income profiles which are allowed to vary with education, after accounting for heterogeneous income growth (and, of course, keeping fixed permanent income). The remaining controls in the model are maternal age at birth interacted with maternal education, and birth year and gender

of the child. Therefore, we extend equation (3.2) to include a large set of controls ( $Z$ ):

$$Y_i = m(PI_i, I_{i2}, I_{i3} | Z_i) + Z_i\delta + \varepsilon_i \quad (3.3)$$

Our argument is that  $I_2$  and  $I_3$  are uncorrelated with  $\varepsilon_i$  after conditioning on  $PI_i$  and all the controls just mentioned. One implication of this argument is that pre-birth investments should be uncorrelated with the timing of income, but may still affect child outcomes. We test this by examining the relationship between having a low birth weight baby and the subsequent timing of parental income. We show below that, although low birth weight is strongly correlated with  $P_i$ , it is uncorrelated with  $I_2$  and  $I_3$ . In addition, in order to control for heterogeneous human capital life-cycle profiles of parents that can interact with the timing of income, instead of using the income measures described above we construct the income residuals from a regression of income on age-education dummies and father fixed effects, and we estimate the role of the timing of income residuals by including these residuals in equation (3.3), instead of  $PI_i$ ,  $I_{i2}$  and  $I_{i3}$ .

Finally, the household characteristics  $Z_i$  can affect the outcomes directly or by interacting with the timing of income. Therefore, we estimate semi-parametrically the function  $m(PI_i, I_{i2}, I_{i3})$  for different groups of household characteristics. For example, we divide the sample in above/below permanent income, above/below median age of the father and median age of the mother, above/below parental years of education and so on. The results are robust to this exercise.

### 3.3.2 Multivariate Local Linear Regression

Equation (3.3) is a partially linear regression model. We adopt a two step method for estimating this model. In the first step we estimate  $\delta$  (the coefficients on  $Z$ ) by using a series approximation for  $m(PI_i, I_{i2}, I_{i3})$ .<sup>5</sup> In the second step we estimate a local linear regression of  $Y_i - Z_i\hat{\delta}$  on  $(PI_i, I_{i2}, I_{i3})$ .

We follow Ruppert and Wand (1994) and to define the multivariate local linear regression estimator. Let  $I_i = (PI_i, I_{i2}, I_{i3})$ . Our goal is to estimate the conditional mean function  $m(I_i) = E(Y|I_i = \mathbf{x})$  for a vector  $\mathbf{x}$ , where  $i = 1, \dots, n$ . The solution

<sup>5</sup>In particular, we approximate  $m(PI_i, I_{i2}, I_{i3})$  as:

$$\begin{aligned} m(PI_i, I_{i2}, I_{i3}) = & \alpha_0 + \alpha_1 PI_i + \alpha_2 I_{i2} + \alpha_3 I_{i3} + \alpha_4 PI_i^2 + \alpha_5 I_{i2}^2 + \alpha_6 I_{i3}^2 \\ & + \alpha_7 PI_i^3 + \alpha_8 I_{i2}^3 + \alpha_9 I_{i3}^3 + \alpha_{10} PI_i I_{i2} + \alpha_{11} PI_i I_{i3} + \alpha_{12} I_{i2} I_{i3} \\ & + \alpha_{13} PI_i^2 I_{i2} + \alpha_{14} PI_i^2 I_{i3} + \alpha_{15} I_{i2}^2 I_{i3} + \dots \end{aligned}$$

(include all two-way and three-way interactions between  $(PI_i, I_{i2}, I_{i3}, PI_i^2, I_{i2}^2, I_{i3}^2, PI_i^3, I_{i2}^3, I_{i3}^3)$ ). Then we can estimate equation (3.3) by least squares.



is the value which minimizes the weighted least squares objective function

$$\sum_{i=1}^n \{Y_i - \alpha - (I_i - x)\beta\}^2 K_{\mathbf{H}}(I_i - \mathbf{x})$$

where  $\mathbf{H}$  is a  $3 \times 3$  diagonal bandwidth matrix and  $K(\cdot)$  is defined as the 3-dimensional product of a univariate uniform kernel function:

$$K(s) = \begin{cases} 0.5 & \text{if } |s| < 1 \\ 0 & \text{otherwise} \end{cases}$$

where  $s = \frac{I_i - \mathbf{x}}{h}$  and  $h$  is the bandwidth.

This results in the estimator for each  $\mathbf{x}$

$$\hat{\alpha} = e^T (I_x^T W_x I_x)^{-1} I_x^T W_x Y$$

where  $e^T$  is the vector with 1 in the first entry and 0 in all others and  $W_x$  is the weighting function at the point  $\mathbf{x}$ .

The choice of kernel is not important for the asymptotic properties of the estimator, as long as it is chosen to be a symmetric, unimodal density, such as the uniform kernel. However, there exists a trade-off in the choice of the number of observations entering the local kernel regressions, determined by the bandwidth  $h$ . A larger bandwidth increases the bias of the estimate but reduces the variance. We expect that  $h \rightarrow 0$  as  $n \rightarrow \infty$ .

We use the following formula to choose our bandwidth, for each covariate:

$$h_j = C * 2 * \sigma_{x_j} h^{\frac{-1}{7}} \tag{3.4}$$

where  $C$  denotes a constant and  $\sigma_{x_j}$  the standard error of component  $j$  of vector  $I$ . We allow  $C$  to vary between 0.5 and 4, in order to examine the robustness of our results to the choice of bandwidth.

Finally, we calculate the standard errors using the formula from Ruppert and Wand (1994).

$$var \{\hat{m}(x, \mathbf{H}) | I_1, \dots, I_n\} = \left\{ n^{-1} |\mathbf{H}|^{\frac{-1}{2}} R(K) / f(x) \right\} v(x) \{1 + o_p(1)\}$$

where  $R(K) = \int K_H(s)^2 ds$ ,  $f(x)$  denotes the conditional density of  $\mathbf{x}$  and  $v(x) = Var(Y | I = \mathbf{x})$  denotes the conditional variance of the outcome. We estimate the

conditional density and variance as follows:

$$\widehat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n \frac{1}{h_1 h_2 h_3} K \left( \frac{I_{i1} - x_1}{h_1}, \frac{I_{i2} - x_2}{h_2}, \frac{I_{i3} - x_3}{h_3} \right)$$

$$\widehat{v}(x) = e^T (I_x^T W_x I_x)^{-1} I_x^T W_x \hat{\epsilon}^2$$

where  $\hat{\epsilon}^2 = Y_i - \widehat{m}(x)$

## 3.4 Results

### 3.4.1 Descriptive Statistics

The descriptive statistics for the sample are reported in Figure 3.1. There are 522,490 child level observations, which are all individuals born in Norway between 1971 and 1980 for whom we were able to collect paternal income data, plus those born in 1986. The average permanent income of the father (in the period between the ages of 0 and 17 of the child) is about £306,100. There is substantial income dispersion (the standard deviation is £116,900). Income in each period (1, 2, and 3) falls with the age of the child because of discounting (we discount all incomes to age 0).

Mothers have on average 11.14 years of schooling, which is slightly lower than the average years of education of the fathers (11.45). Mothers are much younger than fathers at birth (26 vs 29 years of age).

The average years of education of the children in our sample is 12.73. 21% of children drop out from high school, but 39% attend college. The average annual earnings of these children at age 30 is £19,930. As noted above, IQ is only available for males and takes values on a 9 point scale, with a sample average of 5.25, and a standard deviation of 1.79. The average health score for the males is 8.44, indicating that the majority of children achieve perfect physical health on this scale (which has a maximum score of 9). Teen pregnancies occur for 8% of the females in our sample. Finally, the cohort of children for whom we have 10th grade exam information have an average score of 42.75 (out of 60) in all exams combined, and 14.71 (out of 24) in the core exams of Norwegian, English and Mathematics.

	N	Mean	Standard deviation
Paternal Income Period 1, Age 0-5	522490	11.83	4.23
Paternal Income Period 2, Age 6-11	522490	10.25	4.41
Paternal Income Period 3, Age 12-17	522490	8.53	4.68
Paternal Permanent Income, Aged 0-17	522490	30.61	11.69
Mother Years of Schooling	522490	11.14	2.71
Father Years of Schooling	522490	11.45	3.02
Mother Age at Birth	522490	26.26	5.03
Father Age at Birth	522490	29.02	5.74
Child Year of Birth	522490	1975.29	2.88
Years of Schooling	520752	12.73	2.41
High School Drop Out	522490	0.21	0.41
College Attendance	522490	0.39	0.49
Log Earnings age 30	307776	9.90	0.81
IQ (males only)	248801	5.25	1.79
Health (males only)	265959	8.44	1.52
Teenage Pregnancy (females only)	249540	0.08	0.28
Grades	48384	42.75	10.62
Grades Core Subjects	48384	14.71	4.24

Figure 3.1: Descriptive Statistics

### 3.4.2 Parametric Estimates

We first present basic patterns from parametric regressions which then are compared to the semi-parametric estimates. It is particularly interesting to start with a simple version of equation (3.2) where we ignore the timing of income, and consider only the relationship between an outcome of the child,  $Y$ , and the permanent income of the father,  $PI$ . Although it is common to estimate linear models, we will allow the relationship between  $Y$  and  $PI$  to be more flexible. Instead of including  $PI$  linearly in the model, we construct indicator variables,  $q_{PI,i}^k$ , that take value 1 if the paternal income of individual  $i$  is in percentile  $k$  of the distribution of  $PI$  in the sample, with  $k = 1, \dots, 100$ :

$$Y_i = \sum_{k=1}^{100} \phi_{PI}^k q_{PI,i}^k + \varepsilon_i \quad (3.5)$$

The empirical results in this paper will be presented through a series of graphs. We start by focusing on years of education as the outcome of interest. Figure 3.2 plots the relationship between years of education of the child and paternal income constructed from the estimates of equation (3.5). The estimated function is monotonically increasing (except at the very high end) and concave. Increasing paternal

permanent income from £200,000 to £300,000 translates roughly into an increase of 0.5 years of schooling for the child. Figure 3.3 plots the relationship between earnings by age 30 and permanent income. The same increase in permanent income translates into an increase of 10% in earnings of the child by age 30.

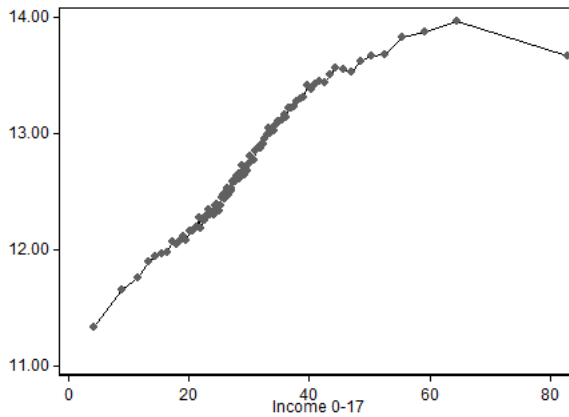


Figure 3.2: Years of schooling v/s paternal Permanent Income.

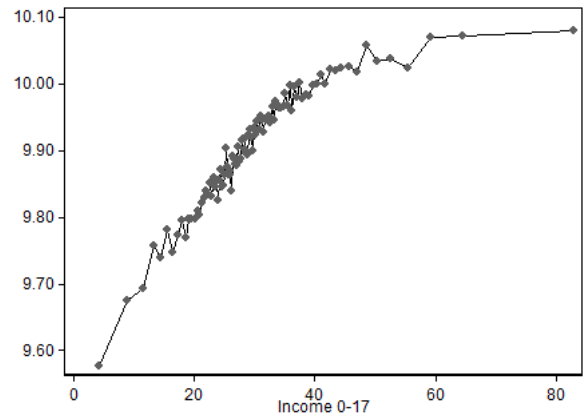


Figure 3.3: Earnings by age 30 v/s paternal Permanent Income

Note: Graphs plot individual coefficients from regression of decile bins for Permanent Income upon child human capital. Income in 2000 prices, £ 10,000s.

Panel B.1 in the Appendix B plots the estimates of equation (3.5) for each of the remaining outcomes. High school dropout rates are declining with  $PI$  for values of  $PI$  below £400,000, and flat after that (remarkably, not going much below 10%). College attendance rates increase substantially throughout the distribution of  $PI$ , and so do IQ scores. Log earnings at age 30 rise steeply with  $PI$  for values of  $PI$  below £400,000, and much more slowly after that. This pattern is somewhat similar to the one we find for high school dropout rates, and curiously, for teenage pregnancy as well. Estimates for the health index are erratic but roughly display an increasing pattern with  $PI$ . All these panels show patterns as expected, and the magnitudes of the relationships between the different outcomes and  $PI$  are very substantial.

The following model introduces the incomes in two periods of childhood (leaving a third out of the model, because of collinearity). One flexible parametric approach of equation (3.3) is the one in which the function  $m(PI_i, I_{i2}, I_{i3})$  is separable in its three arguments:  $m(PI_i, I_{i2}, I_{i3}) = m^1(PI_i) + m^2(I_{i2}) + m^3(I_{i3})$ . The sub-functions  $m^j(PI_i)$  are approximated using dummies for each percentile of the distribution of

$PI$ . In other words, we estimate the following model:

$$Y_i = \sum_{k_1=1}^{100} \phi_{PI}^{k_1} q_{PI,i}^{k_1} + \sum_{k_2=1}^{100} \phi_{I_2}^{k_2} q_{I_2,i}^{k_2} + \sum_{k_3=1}^{100} \phi_{I_3}^{k_3} q_{I_3,i}^{k_3} + Z_i \delta + \varepsilon_i \quad (3.6)$$

where  $q_{PI,i}^{k_1}$  is an indicator that takes the value 1 if the father of child  $i$  has permanent income in percentile  $k_1$  of the distribution of  $PI$  and 0 otherwise.  $q_{I_2,i}^{k_2}$  and  $q_{I_3,i}^{k_3}$  are defined analogously.

We now turn to the main empirical finding of this research, looking at the effects of the timing of income on outcomes. Figure 3.4 and Figure 3.5 relate years of schooling, and the timing of income, represented respectively by  $m^2(I_{i2})$ , and  $m^3(I_{i3})$ . We plot  $m^2(I_{i2})$  keeping the values of all other variables in the model fixed at their means, and analogously for  $m^3(I_{i3})$ . Since the model is separable, each of these functions shows us the partial derivative of the outcome with respect to income in each period, keeping all else fixed.

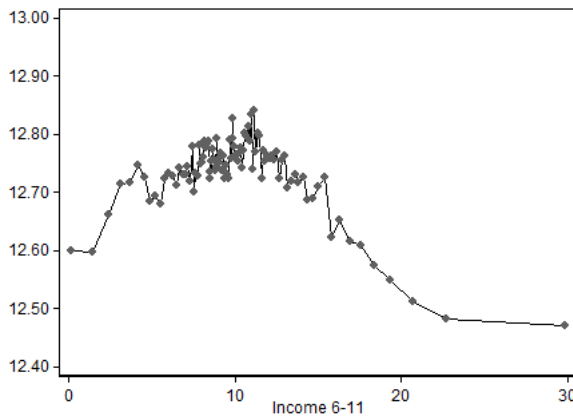


Figure 3.4: Years of schooling v/s paternal Income age 6-11

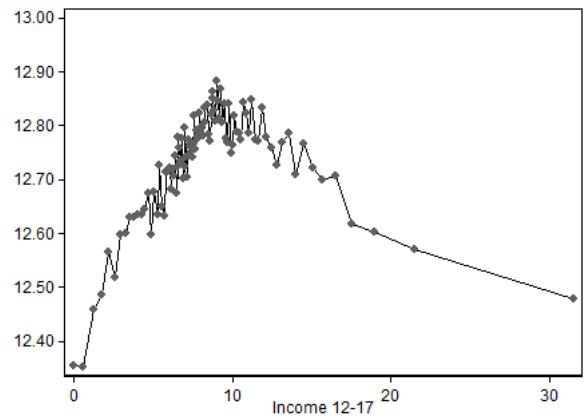


Figure 3.5: Earnings by age 30 v/s paternal Income age 12-17

Note: Graphs plot individual coefficients from regression of decile bins for Permanent Income upon child human capital. Income in 2000 prices, £ 10,000s.

Both  $m^2(I_{i2})$ , and  $m^3(I_{i3})$  are inverse U-shaped. This means that education is maximized when there is some balance between paternal income across periods. It is not desirable (in terms of schooling attainment) to have all father's income concentrated in one period of childhood, regardless of whether it is ages 0-5, 6-11, or 12-17. Below we discuss why this might be the case. Not surprisingly, we have similar findings when we use high school dropout or college attendance as the

outcomes (Panel B.2 and Panel B.3 in the Appendix B). We can reject the hypothesis that  $m^2(I_{i2})$  ( $m^3(I_{i3})$ ) is flat. In particular, we test whether the coefficients  $\phi_{I_2}^{k_2}$  ( $\phi_{I_3}^{k_3}$ ) are all equal to each other (across different values of  $k_2$  ( $k_3$ )). The test is robust to when we drop the coefficients  $\phi_{I_2}^{k_2}$  ( $\phi_{I_3}^{k_3}$ ) which are at the extremes, i.e., for very low or very high values of  $k_2$  ( $k_3$ ). The IQ graphs also display an inverse-U shape, both for income at 6-11 and at 12-17, although they have a fairly long increasing section.

Figure 3.4 and Figure 3.5 relate log earnings at age 30 and the timing of income. The shapes of the graphs are quite different. Child earnings are decreasing with income at ages 6-11, which says that shifting money away from the first period and towards the second period of childhood results in lower labor market outcomes for the child. Child earnings are roughly increasing in income at ages 12-17.

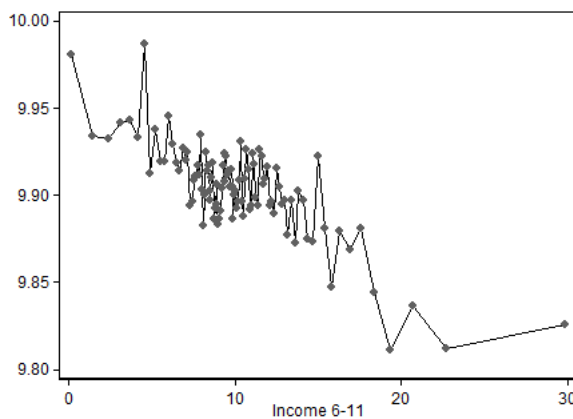


Figure 3.6: Earnings by age 30 v/s Paternal Income age 6-11

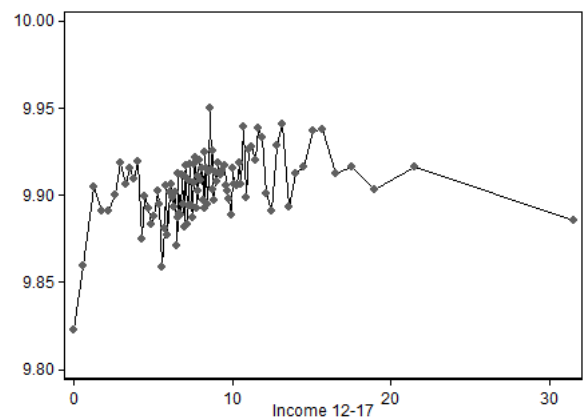


Figure 3.7: Earnings by age 30 v/s Paternal Income age 12-17

Note: Graphs plot individual coefficients from regression of decile bins for Permanent Income upon child human capital. Income in 2000 prices, £ 10,000s.

Panel B.2 and Panel B.3 in the Appendix B also plot the parametric estimates of teenage pregnancy and health. With regards to teenage pregnancy, there is not much of a gradient with income at 6-11, and a pronounced and declining relationship with income at 12-17. This suggests that positive income shocks in the last period of childhood may be particularly important to prevent teenage pregnancy. In terms of self-reported adult health it also seems to be beneficial to shift income towards late childhood.

When we examine grades in high school, it is useful to shift income from ages 6-11 to ages 0-5, indicating that the early years are important. But at the same time it is also important to shift income towards ages 12-17. Notice that, for grades

in high school, we increase the size of the bins over which we evaluate  $(PI_i, I_{i2}, I_{i3})$ . This is because the sample size is so much smaller for this outcome.

Finally, from this section we can conclude that despite the fact that permanent income is more important than the timing of income for all the outcomes, the timing of income is still relevant. We have seen that an increase in permanent income from £200,000 to £300,000 leads to an increase in earnings by age 30 of 10% and 0.5 in years of schooling. But similar shifts of per-period incomes lead to sizable effects. A shift in paternal earnings from age 0-5 to age 6-11 from £100,000 to £200,000 leads to a decrease earnings by age 30 in 5% and to a decrease in 0.2 in years of schooling. A shift in paternal earnings from 0-5 to 12-17 from £100,000 to £200,000 leads to an increase of 1-2% in earnings by age 30 and to a decrease in 0.3 years of schooling.

### 3.4.3 Semi Parametric Estimates

In this section we present semi-parametric estimates of  $m(PI, I_2, I_3)$ , following the method laid out in section 3.3.2. In order to implement it we need to first create a grid of evaluation points for  $m(PI, I_2, I_3)$ , which is tridimensional. We take 19 points for each income variable  $(PI, I_2, I_3)$ , corresponding to the ventiles (1/20) of each variable's distribution. This gives us a tridimensional grid with 6,859 points ( $= 19 * 19 * 19$ ).

It is standard practice to trim the data so to avoid spurious results driven by small cells. Therefore, we drop 2% of observations, corresponding to the cells with the smallest number of observations. In our main results we use a uniform kernel and choose the bandwidth using the formula in equation (3.4), setting  $C = 1$ . Below we show that our results are robust to the choice of kernel and bandwidth.

The estimates of  $m(PI, I_2, I_3)$  are presented through a series of two dimensional graphs, where the y-axis is the outcome of interest, and in x-axis is one of the income variables. The downside of this type of presentation is that we can only vary one of the incomes being considered at a time, which means that we need to fix the remaining two variables (we also fix the remaining control variables at their mean values). Therefore, we need to use multiple figures for each outcome. The advantage of such an apparently cumbersome approach is that the graphs are straightforward to read.

For each outcome, we present three sets of graphs. In the first set, we fix  $PI$  and  $I_3$  at three different values each (the third, fifth, and seventh deciles of the distribution of each variable), and vary only  $I_2$ , for a total of nine possible combinations. These are presented in nine different panels, which plot  $m(PI, I_2, I_3)$  against  $I_2$  (for

given values of  $PI$  and  $I_3$ ). At the top of each panel we display the values at which we are keeping  $PI$  and  $I_3$  fixed.

Since  $PI = I_1 + I_2 + I_3$ , if we keep  $PI$  and  $I_3$  fixed then it is not possible to vary  $I_1$  and  $I_2$  independently. Therefore, as we move towards the right in the x-axis we see how the outcome varies as we shift income from period 1 to period 2. The support of  $I_2$  over which we can evaluate  $m(PI, I_2, I_3)$  is not the same across all panels because there are either infeasible values for  $I_2$ , or values of  $I_2$  which are feasible but for which there are no observations in the sample (for given combinations of  $PI$  and  $I_3$ ). The second set of panels keeps  $PI$  and  $I_2$  fixed, and varies  $I_3$  (so we are shifting income from period 1 to period 3). The third set of panels keeps  $PI$  and  $I_1$  fixed, and varies  $I_2$  (so we are shifting income from period 3 to period 2).

Below each panel we display two other parameters and respective standard errors,  $\alpha_1$  and  $\alpha_2$ . For each panel, let  $H$  be the highest point of support for the income variable being used in that panel, let  $L$  be the lowest point of support, and  $M$  be the median point of support (which would correspond to exactly the 50th percentile of the distribution of that income variable if all graphs had full support). Take the case where we fix  $PI = \overline{PI}$  and  $I_3 = \overline{I_3}$ , and we let  $I_2$  vary. Then we define:

$$\begin{aligned}\alpha_1 &= m(\overline{PI}, M, \overline{I_3}) - m(\overline{PI}, L, \overline{I_3}) \\ \alpha_2 &= m(\overline{PI}, H, \overline{I_3}) - m(\overline{PI}, M, \overline{I_3}).\end{aligned}\tag{3.7}$$

$\alpha_1$  is the difference between the values the outcome takes in the median and lower extreme of the support of  $I_2$ , while  $\alpha_2$  is the difference between the values the outcome takes in the median and upper extreme of the support of  $I_2$ . If  $m(\overline{PI}, I_2, \overline{I_3})$  did not vary with  $I_2$  (in which case the timing of income is irrelevant, at least when we compare first and second period incomes) we would expect  $\alpha_1 = \alpha_2 = 0$ , so these parameters help us quantify the importance of the timing of income.

### 3.4.3.1 Schooling Attainment

We begin by focusing on years of schooling of the child as the outcome of interest. Figure 3.8 plots years of schooling against  $m(\overline{PI}, I_2, \overline{I_3})$ , fixing  $PI$  and  $I_3$  at their median values. The interpretation of this graph is how schooling varies when  $I_1$  is shifted to  $I_2$  at the median values of  $PI$  and  $I_3$ . The solid line (with the dashed standard errors) corresponds to  $m(\overline{PI}, I_2, \overline{I_3})$ . The scale of this line is given on the vertical axis located on the left of the graph. The dotted line which is declining in every panel corresponds to the missing income. In this case, it is equal to  $I_1 =$



$\overline{PI} - \overline{I}_3 - I_2$ . The scale of this line can be read on the vertical axis located at the right of the graph. Figure 3.9 plots years of schooling against  $m(\overline{PI}, \overline{I}_2, I_3)$ , fixing  $PI$  and  $I_2$  at their median values. The interpretation of this graph is how schooling varies when  $I_1$  is shifted to  $I_3$  at the median values of  $PI$  and  $I_2$ . Finally, Figure 3.9 plots years of schooling against  $m(\overline{PI}, \overline{I}_2, I_3)$ , fixing  $PI$  and  $I_2$  at their median values. The interpretation of this graph is how schooling varies when  $I_1$  is shifted to  $I_3$  at the median values of  $PI$  and  $I_2$ .

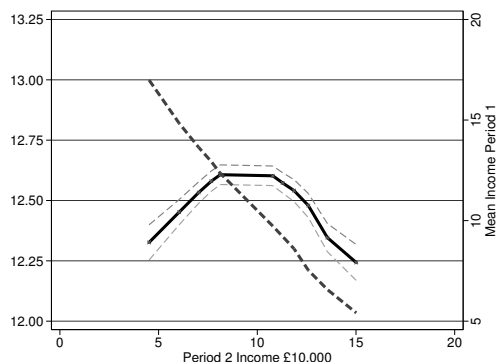


Figure 3.8: Schooling Attainment v/s Paternal Income age 6-11.  $PI$  and Income age 12-17 fixed at the median

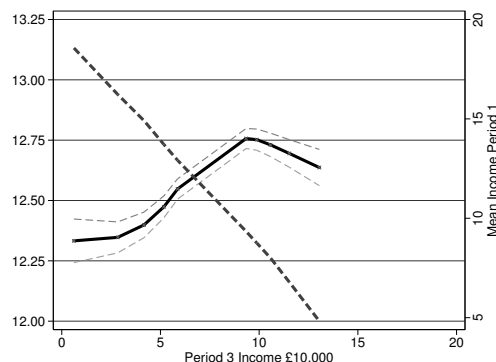


Figure 3.9: Schooling Attainment v/s Paternal Income age 12-17.  $PI$  and Income age 6-11 fixed at the median

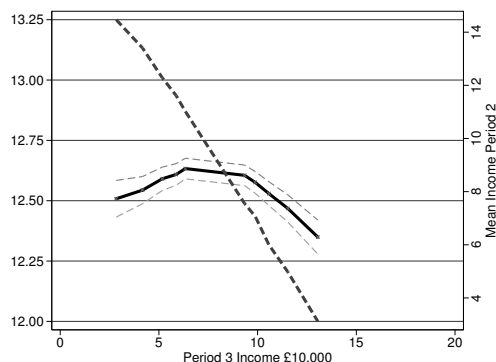


Figure 3.10: Schooling Attainment v/s Paternal Income age 12-17.  $PI$  and Income age 0-5 fixed at the median

Note: Straight lines are the semi-parametric estimates of schooling attainment. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

Remarkably, the figures display an inverse U-shape, the same as the parametric plots. We compute  $\alpha_1$  and  $\alpha_2$  for each panel, and we are able to reject that the slope of these functions is equal to zero. What this says is that across different values of  $PI$  and  $I_3$  the years of schooling of the child are maximized when there is some balance between period 1 and period 2 income. If income is too concentrated in either period 1 or period 2, then one can improve education outcomes of children by shifting income towards the other period. The (discounted annual) level of  $I_2$  at which the maximum is achieved is roughly between £8000 and £12000 (a little higher for richer households, and lower for poorer households).

We obtain the same data patterns if we present the same plots for the figures above but fixing permanent income and the per-period income that is left aside at different deciles. Panel B4 shows how years of schooling change with  $I_2$ , relative to  $I_1$ . Panel B5 shows how years of schooling change with  $I_3$ , relative to  $I_1$ , and Panel B6 shows how years of schooling change with  $I_3$ , relative to  $I_2$ . At the top of each panel we display the values at which we keep  $PI$  and the income that is left aside fixed, which are either the third, fifth or seventh deciles of the respective distributions. We compute  $\alpha_1$  and  $\alpha_2$  (from equation (3.7)) for all panels, and we find statistically significant upward and downward slopes for each figure. This indicates that these curves are definitely not flat.

These results imply that the timing of income shocks is relevant for human capital formation. If the timing of income was irrelevant then all these graphs would be horizontal lines, with only permanent income being relevant for human capital outcomes. Most likely, the reason why timing matters is that the timing of income shocks affects the timing of investments in human capital. This will happen if parents have imperfect insurance possibilities against income shocks, and if the technology of skill formation is complementary in investments across different periods of childhood.

In a scenario where parental investments will react to income shocks, the shape of the curves in panel B4 will tell us something about the technology of skill formation. So suppose we compare children in families with very volatile incomes, and therefore, volatile investments, with children in families with stable incomes (between periods 1 and 2). Then the latter will do much better, keeping constant total (permanent) income across the childhood years, if the technology exhibits complementarity between period 1 and period 2 investments, since in that case, you will want to maintain a stable flow of investments over the life of the child. All these mechanisms will be analyzed in light of simulations of a dynamic structural model

of parental investments in children with a technology of skill formation, extensively discussed in 3.6.

Panel B5 examines the trade-offs between periods 1 and 3 income (keeping fixed period 2 income and permanent income) and show a similar inverse-U shape, although it is less pronounced than in the previous figures. Some of the graphs display curves that are mainly monotonically decreasing. Again, if we take the view that uncertainty and partial insurance cause investments in children to react somewhat to income shocks, these figures are telling us something about the technology. In particular, they are telling us that investments in the early years are particularly productive (or particularly cheap) relative to parental investments in the adolescent years, and that investments very early and very late in the life of the child may be quite substitutable.

Finally, Panel B6 examines trade-offs between periods 2 and 3 income (keeping fixed period 3 income and permanent income). Most of the figures still display an inverse-U shape, although a few indicate that it is better to delay income from period 3 to period 2.

### 3.4.3.2 Log Earnings at age 30

Next we present results for the case where the outcome is log annual earnings at age 30. Some of them are shown in the next figures, but the full set of figures are shown in panels B7-B9 of Appendix B. Keeping  $PI$  and  $I_3$  fixed, shifting income away from period 1 and towards period 2 leads to a sharp reduction in log earnings, followed by a flattening of the relationship. When we shift income from period 1 to 3 log earnings seem to be slightly increasing meaning that late investments are comparatively better rewarded. However, when we shift income from period 2 to 3, we observe similar patterns we had for education.

The slopes of these curves are remarkably steep. For low values of  $I_1$ , a £100000 shift in income from  $I_1$  to  $I_2$  (keeping  $PI$  and  $I_3$  fixed) leads to a 5-15% decline in wages. A shift of either £100000 in  $I_1$  or  $I_2$  towards  $I_3$  generates gains in earnings close to 5%. These figures are very large, especially in light of the fact that a £100000 increase in  $PI$  is associated with a 10% increase in log earnings at age 30.

### 3.4.3.3 High School Drop Out and College Attendance

Instead of years of schooling, it is useful to consider high school dropout rates and college attendance rates separately, since they correspond to two groups of

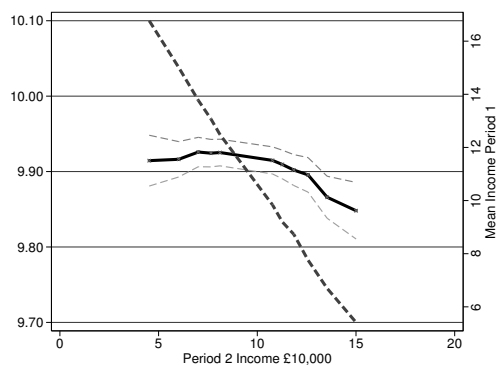


Figure 3.11: Earnings by age 30 v/s Paternal Income age 6-11. *PI* and Income age 12-17 fixed at the median

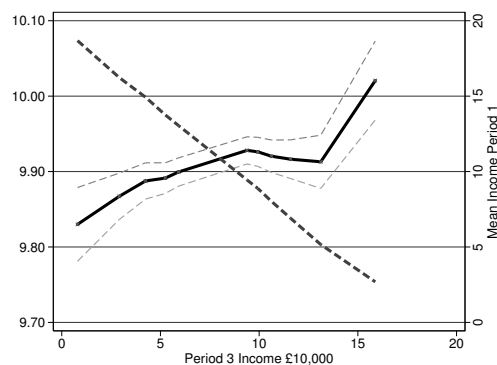


Figure 3.12: Earnings by age 30 v/s Paternal Income age 12-17. *PI* and Income age 6-11 fixed at the median

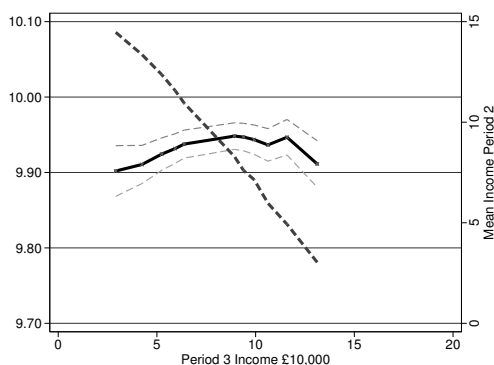


Figure 3.13: Earnings by age 30 v/s Paternal Income age 12-17. *PI* and Income age 0-5 fixed at the median

Note: Straight lines are the semi-parametric estimates of earnings by age 30. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

individuals, one in the lower tail and the other in the upper tail of the education distribution. Here we show just two examples of our results plotting outcomes against period 2 income, fixing permanent income and period 3 income at the median values. However, all the trends are the same as schooling attendance, no matter which income is shifted. In panels B10 and B11 we just plot the data patterns when income is shifted from period 1 to period 2.

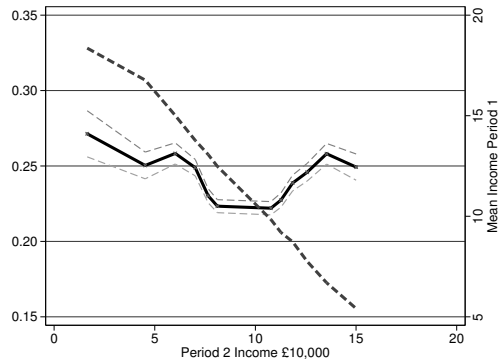


Figure 3.14: High School Dropout rates v/s Paternal Income age 6-11. *PI* and Income age 12-17 fixed at the median

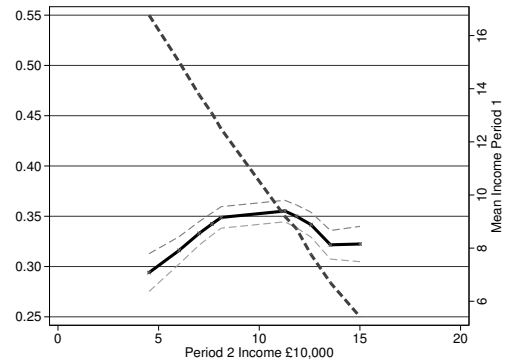


Figure 3.15: College attendance v/s Paternal Income age 6-11. *PI* and Income age 12-17 fixed at the median

Note: Straight lines are the semi-parametric estimates of child's human capital measures. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

Not surprisingly, results are quite similar to the ones we showed for years of education. High school dropout rates are minimized, and college attendance rates are maximized, when incomes are balanced between the early and middle childhood years (periods 1 and 2), keeping permanent income and income in adolescence fixed. When we increase income in adolescence (period 3), educational outcomes appear to improve when this is done at the expense of early childhood but worsen when at the expense of middle childhood income.

Again, we can clearly reject that these curves are flat, by computing both  $\alpha_1$  and  $\alpha_2$  for each panel. The magnitudes of these impacts are substantial. An increase in permanent income of £100000 is associated with roughly a 10% decline in high school dropout rates and a 10% increase in college attendance rates. In comparison, a £100000 shift in income from period 1 to period 2 leads roughly to a 4% decrease in high school dropout and a 6% increase in college attendance.

### 3.4.3.4 Other Outcomes

Panels B12-B14 also summarize some of the semi-parametric estimates of the remaining outcomes. Because of space concerns, we only show data patterns evaluated at median permanent income and per period incomes left aside. Panel B12 shows the trade-off between  $I_1$  and  $I_2$ , Panel B13 the trade-off between  $I_1$  and  $I_3$ , and Panel B14 the trade-off between  $I_2$  and  $I_3$ .

IQ estimates do not deliver an obvious pattern when we study the trade-off between  $I_1$  and  $I_2$ . However, when we study the trade-offs between  $I_1$  and  $I_3$ , and  $I_2$  and  $I_3$ , the results clearly indicate that shifting income towards adolescence is associated with higher and lower levels of IQ respectively.

With regards to health, and in contrast to what we have seen so far, delaying income from  $I_1$  to  $I_2$  seems to lead to poorer health outcomes although the patterns are not clear. We cannot reject that about half the curves are flat lines. The estimates are also relatively more imprecise in this case than for the outcomes studied so far.

Teenage pregnancy is minimized when there is a balance between  $I_1$  and  $I_2$ , and when there is a shift in income from  $I_1$  to  $I_3$ , and a balance between  $I_2$  and  $I_3$ . Nevertheless, because teenage pregnancy is only observed for females, and that it is a relatively infrequent phenomenon, these estimates are more imprecise than the ones presented above for other outcomes.

Unfortunately the nonparametric results for grades in school are too imprecise to be informative.

## 3.5 Tackling Endogeneity

In our main empirical specification we can't really argue that the error term is orthogonal to the timing of income, even after controlling for permanent income and household characteristics. The true production function driving final outcomes is a function of investments, which are unobserved, and households jointly decide investments, consumption and labor supply according to many factors like preferences, the underlying technology or budget constraints. Therefore, it is likely that the timing of income is a function of labor supply, correlated to investments. We focus on three reasons why this can happen. First, if the household temporarily changes labor supply and the timing of income is a choice. One clear example of this is when family income in the early years of childhood adjusts in response to maternity leave choices. If the mother chooses to take a prolonged period of leave,

paternal income may increase in response. Second, the timing of income is also endogenous if age-earnings profiles are heterogeneous. Particularly, income paths can be front-loaded or backloaded depending on the level of human capital. For example, parents who are high-school degree are likely to start earlier in the labor market but their income profiles are less steeper than college graduates, who start later. Third, income profiles can be endogenous to family-specific unobserved preferences over child-rearing. Below we discuss each of these issues using the human capital outcome years of education.

### 3.5.1 Maternity leave choices

The first strategy already implemented in the previous estimates is to use only father's incomes, which are usually more stable in the labor market. However, we cannot rule out that paternal labor supply choices are also affected by child-rearing activities. In Figure 3.16 we show one of the results of re-estimating the model excluding income in years 0-2 when schooling is plotted against income in period 2 at median values of permanent income and income in period 3. The effect of the timing of income is estimated now defining income in the three periods as 3-7, 8-12 and 13-17 with permanent income defined as the sum across these periods. The results look very similar, showing if anything more definition in the curvature of the inverse u-shape relationship. Figures 3.17 and 3.18 show the same idea but shifting income from periods 1 to 3 and 2 to 3 respectively. The inverse u-shape remains statistically significant.

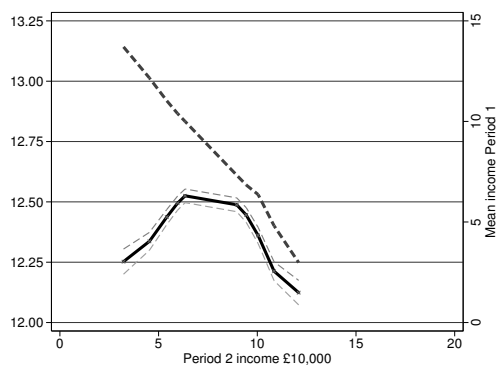


Figure 3.16: Schooling attainment v/s Paternal Income age 6-11 excluding age 0-2, *PI* and Income age 12-17 fixed at median values

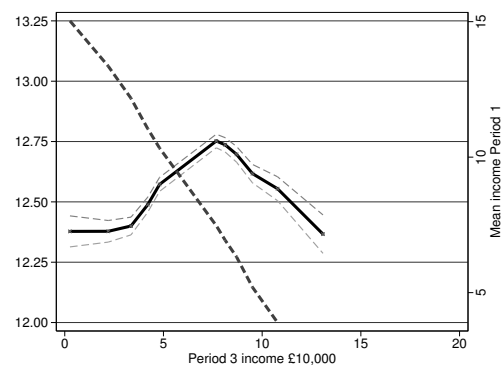


Figure 3.17: Schooling attainment v/s Paternal Income age 12-17 excluding age 0-2, *PI* and Income age 6-11 fixed at median values

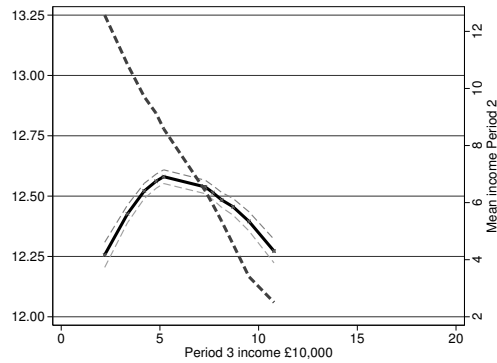


Figure 3.18: Schooling attainment v/s Paternal Income age 12-17 excluding age 0-2, *PI* and Income age 0-5 fixed at median values

Note: Straight lines are the semi-parametric estimates of child's human capital measures. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

### 3.5.2 Heterogeneous age-earning profiles

To address these concerns we adopt three strategies. First, it could be that paternal incomes are front-loaded or backloaded as a consequence of heterogeneous human capital. We estimate the same model using schooling attainment as outcome but controlling for father and mother education, interacted with father and mother's age at birth. The set of plots (not shown) look very similar to those discussed above, and the inverse U-shape is robust and significant. Second, it could be that heterogeneous income profiles are driven by unobserved individual-specific life-time income trends. We address this concern by reconstructing income profile of fathers calculating the difference between income post-childhood (using age 18-20) and pre-childhood (three years prior to birth). We then use as per period incomes the deviations from individual trends. Again, we draw the same conclusions as in the main specification. Finally, the remaining endogeneity may take the form of differential variance of income. It could be that parents with less volatile income profiles are "better" parents than those with fluctuating income profiles. We control for this by running a fixed effect regression of fathers' income on dummy variables for education and age. We then calculate for each individual the error term in each period and take the variance of this as an additional control in the main semi-parametric regressions.



### 3.5.3 Heterogeneous preferences

We test our semi-parametric estimations against heterogeneous preferences. In particular, we tackle unobserved family preferences towards child-rearing potentially correlated to the timing of income by exploiting within family variation. The idea is to capture differential effects of the timing of income in human capital across siblings of different ages. Figures 3.19 and 3.20 show the estimations of equation 3.3, showing that our results are robust.

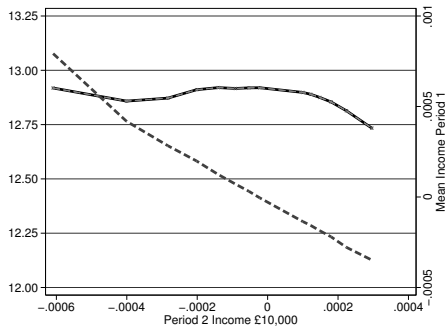


Figure 3.19: Schooling attainment v/s Paternal Income age 6-11 controlling for Fixed Effects, *PI* and Income age 12-17 fixed at median values

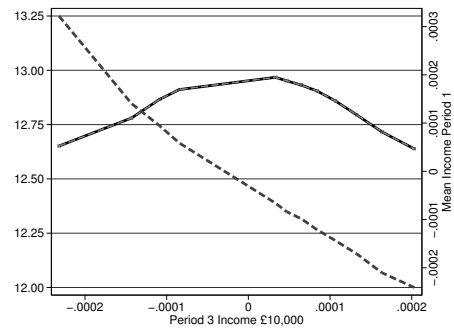


Figure 3.20: Schooling attainment v/s Paternal Income age 12-17 controlling for Fixed Effects, *PI* and Income age 6-11 fixed at median values

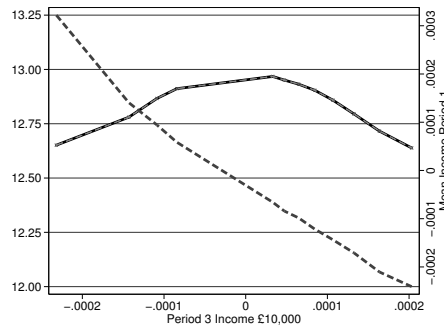


Figure 3.21: Schooling attainment v/s Paternal Income age 12-17 controlling for Fixed Effects, *PI* and Income age 0-5 fixed at median values

Note: Straight lines are the semi-parametric estimates of schooling attainment. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

### 3.5.4 Further Robustness checks

We subsequently control for additional sources of bias that could potentially drive our results. First, as discussed in section 3.2, we have used a fixed discount rate to construct our measures of per period income and permanent income. To address concerns that the chosen discount rate is not appropriate for our panel we repeat estimations using different fixed rates (0%,2%,4%,6%,10%,15%), and we also re-estimate the model using time-varying discount rates using official real interest rates data for Norway between 1971 and 1998. In all the cases the shapes are preserved. Second, we control for household composition variables like marital status and the number of children. These are likely to interact with the relationship between the timing of family income and child outcomes, but are endogenous to family income and therefore excluded from the specification. Third, we test for the robustness of the results to bandwidth choice, firstly reducing and then increasing the bandwidth. Using equation (4), bandwidth in the main paper is defined by setting  $C$  equal to 2. We varied this by re-estimating using the smaller bandwidth defined by  $C = 1$  and the larger bandwidth defined by setting  $C = 3$ . Results using the smaller bandwidth are more noisy, but the general patterns remain. Finally, in the main specification we only take income from biological fathers, irrespective of marital break up and further family formation. This may lead to problems if for example mothers re-marry, in which case income from the non-biological father becomes the main income source. We re-estimate our model selecting only families which do not experience marital break-up. The patterns remain remarkably similar.

Finally, one implication of our assumptions is that income fluctuations should not predict pre-birth investments, unless they are related with permanent traits of parents which also would have independent effects on all the outcomes we consider. We test this using an indicator for low birth weight, which is strongly correlated with our permanent income measure. We find that the timing of income fluctuations does not predict whether a child is low birth weight. Figures 3.22 and 3.23 present results from a regression of a low birthweight dummy on  $(PI, I_2, I_3)$ . It is not possible to reject that these lines are flat, suggesting that our methodology is indeed valid.

As a summary, the conclusions of our paper are robust to a range of checks for our identification strategy and specification of the human capital equation. Where we do find deviations, it tends to be in the direction that the robustness checks find stronger evidence of inverse u-shaped relationships between the timing of income and child education.

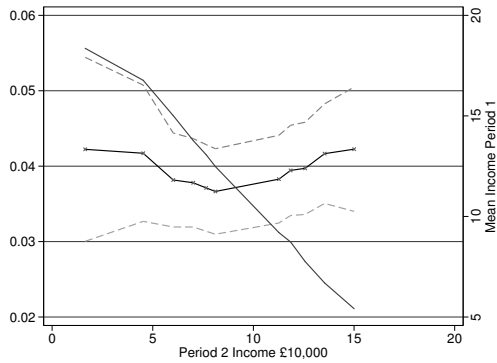


Figure 3.22: Low birth weight v/s Paternal Income age 6-11, *PI* and Income age 12-17 fixed at median values

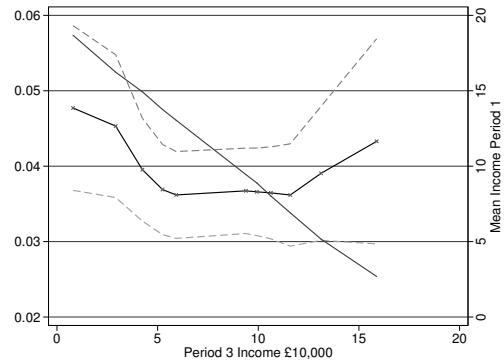


Figure 3.23: Low birth weight v/s Paternal Income age 12-17, *PI* and Income age 0-5 fixed at median values

Note: Straight lines are the semi-parametric estimates of low birth rates. Dashed line is Mean Income of the period of reference. 95% confidence intervals shown. Income in 2000 prices, £ 10,000s. Estimates control for dummies for paternal education interacted with age and maternal education interacted with age, paternal income profile, gender and child year of birth.

### 3.6 The Mechanisms

Why the timing of income drives human capital formation of children when they become adult after controlling for permanent income? In this section we attempt to address this question by estimating dynamic behavioral models of parental investments that allow us to understand the inverse U-shapes found in different schooling outcomes. For example, one can think that the downward sloping section is driven by credit constraints. If the technology of skill formation is complementary in investments across periods and credit constrained parents cannot borrow enough from the future the level of investments during the early childhood would be sub-optimal. It is also interesting that there is an upward sloping section in each curve. One would think that, for a given level of permanent income, it should not be worse to receive all of the income in the first period than to receive it in spread out payments over different periods of childhood. If permanent income is fully available at time zero then one can allocate it freely across periods just by saving the appropriate amount, regardless of whether or not one can borrow. However, this reasoning ignores that parents face multiple sources of uncertainty when they make investment decisions. One of them is income uncertainty. When faced with income shocks, parents change their investments in children, unless they have perfect insurance. Savings alone cannot provide perfect insurance. Another one is uncertainty about

the ability of the child or about the technology of skill formation. If parents are not sure about the ability of the child, parents may want to postpone investment until more of this uncertainty is revealed (see Altonji, Hayashi, Kotlikoff, Journal, and December (1997)).

Furthermore, there could be issues related to preferences. For example, the parents' objective function may include other arguments beyond child's schooling, so depending on how the marginal rate of substitution between parental consumption and investments in children changes with the child's age, delayed parental income could lead to higher investments in children. We discuss these ideas in detail in this section, where we simulate different models of parental investments in children, and examine their implications for the impact of the timing of income shocks on human capital formation.

### 3.6.1 A simple model

The true production function of child outcomes when they become adults is a complex function of parental investments in children across different childhood stages. Furthermore, we know from the literature of the economics of child development that the timing of parental investments in children matter for the process of skill formation if investments have some degree of complementarities across ages <sup>6</sup>. There is a wide evidence from this literature that different types of skills are formed in sensitive and critical periods, and therefore parental investments are dynamic complementary. This means that under-investments in early childhood are harder to be remediated with later investments, in the same way that early investments need to be followed-up in the next childhood stages, so human capital is maximized when investments across periods are balanced.

We realize that if the timing of income matters for human capital formation is because income shocks are likely to be correlated with parental investments. With complete markets and no uncertainty, there should be no differential effect of the timing of income and the slopes of our two-dimensional graphs should be horizontal. Parents can borrow and save to smooth consumption and their investment decisions in child human capital. Consequently, for a given level of permanent income, shifting income across periods of childhood should have no effect upon child human capital.

Unfortunately, in the data we do not observe parental investment directly, but parental income, so we support our data findings building in theoretical models

---

<sup>6</sup>Examples are Cunha and Heckman (2006), Cunha, Heckman, and Schennach (2010), Cunha and Heckman (2009)

of parental investments in children. In these models, parents decide consumption, savings and child investments in each period of the childhood according to a budget constraint which includes shocks to income and borrowing constraints. Parental investments in every period drive child's human capital accumulation entering as inputs of a technology of skill formation.

We perform a simple estimation exercise of such models using the data of schooling attainment to investigate the extent to which complementary investments are responsible for schooling being maximized when the timing of income is balanced. Two arguments emerge from these models. First, if the technology of skill formation is complementary in investments, credit constrained parents would like to invest more and borrow against their permanent income but they can't, so they become reactive to income shocks. If these credit constraints were released, then shifting disposable income from the future to the present could explain the decreasing part of our inverse U-shaped curves. Second, for now we investigate the role of income uncertainty explaining the upward slopes of our graphs. Parents that are risk averse and face uncertainty about their future income may decide not to invest optimally and over-save in the presence of positive income shocks.

Consistently with the empirical strategy, we aim at estimating the technology of skill formation for a four-period model, three periods for the childhood stages and a terminal period for the realization of the child's human capital when she becomes an adult. Preference parameters and the income process are calibrated so that reproducing the income support found in the data. Uncertainty is incorporated through income shocks, which requires a recursive numerical solution using dynamic programming techniques.

## The technology of skill formation

Cunha, Heckman and Schennach (2010) estimate a technology of skill formation with US data proposing a cobb-douglas functional form. In our setup, we use a CES production function, a more flexible specification that do not impose an ex-ante level of substitutability across investments. In this specification, human capital at age  $(t + 1)$ ,  $(h_{t+1})$ , is a function of the current human capital stock  $(h_t)$  and the investment  $(x_t)$

$$h_{t+1} = \delta \left[ \gamma_t h_t^\phi + (1 - \gamma_t) x_t^\phi \right]^{\frac{\rho}{\phi}} \quad (3.8)$$

Under this specification,  $(\gamma_t)$  is the self-productivity of the current stock of human capital in  $(t)$  as opposed to the productivity of new investments,  $(\rho)$  denotes

the degree of concavity of the production function, and  $(\phi)$  the degree of complementarity of investments across periods, where  $(1/1 - \phi)$  is the elasticity of substitution. If investments are perfect substitutes then  $(\phi = 1)$  and  $(\phi = -\infty)$  if investments are perfect compliments. Investments are said to be dynamic complementary if the change in returns to early (late) investment by increasing the level of later (early) investment is positive, that is, if  $\frac{\delta^2 h_t}{\delta x_i \delta x_j} > 0$  and viceversa. Finally,  $(\delta)$  is a scale parameter that anchors unobserved skills into observable outcome measures (years of schooling).

We aim at estimating this flexible technology calibrating the rest of the parameters for preferences. For robustness, we try estimation changing the rest of calibrated parameters one at a time and we always find that the technology is complementary in investments.

## Preferences

We impose the assumption that households are indivisible decision-maker units with standard CRRA preferences of the form

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma} \quad (3.9)$$

where  $(c_t)$  is per-period household consumption. Under this specification, households can react to income uncertainty over-saving if they face positive income shocks. Accumulated assets are described by the standard budget constraint

$$a_{t+1} = (1+r)a_t + y(\varepsilon_t) - px_t + c_t \quad (3.10)$$

where  $(p)$  is the relative price between investment and consumption,  $(a_t)$  is the current stock of assets, and  $(y(\varepsilon_t))$  is the per-period income, which is a function of iid income shocks  $(\varepsilon_t)$ , minimum income  $(w_{min})$ , and a flat wage  $(w)$ .

We add further heterogeneity in incomes to match the support of schooling and per-period incomes found in the data by including three unobserved types. We split the sample into three groups of equal size and we estimate different wage slopes and initial endowments by type. Denoting the types by  $(k)$  and the type-specific wage slope by  $(w_k)$ , the income process is defined by

$$y_k(\varepsilon_t) = w_{min} + w_k \exp(\varepsilon_t) \quad (3.11)$$

As we find that different types have significant differences in average years of schooling, we run regressions of wage growth on average schooling to obtain estimates of differences in years of schooling across types. We use estimated differences w.r.t to type 1 as the initial endowment for the underlying process of skill accumulation ( $H_{0,k}$ ).

### The household problem

As a matter of simplification, we model consumption and savings at the household level and each household has only one child. The recursive formulation then becomes,

$$V_t(h_t, a_t) = \text{Max}_{c_t, i_t} \left\{ \frac{c_t^{1-\sigma}}{1-\sigma} + \beta \int_{\varepsilon_{t+1}} [V_{t+1}(h', a' | \varepsilon')] d\varepsilon \right\}$$

*s.to.*

$$c_t + pi_t + a_{t+1} = y_t + (1+r)a_t$$

$$y_{t,k} = w_{min} + w_k e^{\varepsilon_t}$$

$$h_{t+1} = \delta \left[ \gamma_t h_t^\phi + (1-\gamma_t) i_t^\phi \right]^{\rho/\phi} h_0$$

$$c_t, i_t, a_T \geq 0$$

$$a_t \geq -a$$

$$V_T(h_T, a_T, \varepsilon_T) = \eta \frac{(h_T)^{1-\sigma}}{1-\sigma} + \tau [(1 - \exp(-a_T))]$$

We solve the model incorporating different parameters for borrowing constraints. The results we show below relate to  $a = 0$  (no credit capacity). At each childhood stage parents schooling of their children, assets and income shocks, and decide consumption and investments taking into account expected incomes. The state space is given by ( $\Omega_t = \{h_t, a_t\}$ ) and the value function ( $V(\Omega_t)$ ).

The model is solved by backward recursion using numerical methods, where parents maximize the bellman equation subject to the budget constraint, the technology of skill formation, and a parametrized terminal value function in ( $T = 4$ ), in which parents value both the level of human capital acquired by their children when they become adults ( $h_T$ ), and household assets ( $a_T$ ).

Table 3.1 shows the set of calibrated parameters for estimation.

Parameter	Value
Risk Aversion ( $\sigma$ )	0.5
Discount factor ( $\beta$ )	0.96
Interest rate ( $r$ )	0.5
Relative price investment/consumption ( $p$ )	1
Parental valuation of child's total years of schooling ( $\eta$ )	12
Parental valuation of assets when children become adults ( $\tau$ )	12
Minimum wage ( $w_{\min}$ )	1
Wage slopes by type ( $w_k$ )	{7,7.65,11}
Variance of income shocks ( $\sigma_\varepsilon^2$ )	0.1
Initial endowment ( $H_{0,k}$ )	{1,1,0282,1.0483}

Table 3.1: Calibrated Preference Parameters

### 3.6.2 Solution and Estimation

The model is solved by backward recursion integrating the value function in the next period ( $V_{t+1}$ ) over income shocks by using standard Gauss-Hermite quadrature methods. In each period, optimal investments and consumption involves the approximation of the expected value function for each possible combination of assets and human capital, which evolve according to the optimal decisions and the equations governing dynamics (the budget constraint and the technology of skill formation). In order to handle a large state space, we approximate the expected value function ( $EV_{t+1}$ ) by using Cheyshev Polynomials. The chebyshev coefficients by each period are used to evaluate the Bellman equation and solve for the optimal policy functions of investments and savings ( $i(h_t, a_t, \varepsilon_t)$  and  $a'(h_t, a_t, \varepsilon_t)$ ). We use these policy functions to simulate data from 100,000 random draws of income shocks  $\varepsilon_t$  for estimation. The details of approximation procedure are shown in the Appendix B.

We estimate the model by Simulated Method of Moments (Gourieroux, Monfort, and Renault (1993)). The set of data moments is  $\alpha$ , and for a given value of the structural parameters  $\theta = \{\phi, \gamma_t, \delta\}$ , the model is used to simulate the same set of moments  $\alpha^s(\theta)$  so that minimizing

$$\text{Min}_\theta(\alpha - \alpha^s(\theta))W(\alpha - \alpha^s(\theta))$$

where  $W$  is a weighting matrix.

By using our policy functions derived from the solution, we simulate a large dataset incomes and human capital  $\{I_1, I_2, I_3, PI, H\}$  which is then used to estimate the same Local Linear Regressions used with the actual data, and to construct the



analogous two-dimensional graphs shifting incomes from period 1 to 2 at fixed deciles of  $(I_3)$  and  $(PI)$  and shifting income from period 1 to 3 at fixed deciles of  $(I_2)$  and  $(PI)$ . We match the mean value of years of schooling for five percentiles 10, 25, 50, 75 and 90 of income in period 2 (shifted income from period 1 to 2), and the same five percentiles of income in period 3 (shifted income from period 1 to 3). We use as a weighting matrix the diagonal of the optimal weighting matrix, which is computed using the standard errors of the income percentiles from the data.

$$W = \text{diag}(VCV(\alpha)^{-1})$$

### 3.6.3 Results

Our estimation results of the technology of skill formation are shown in Table 3.2. Self-productivity parameters  $\gamma_t$  show that the current stock of human capital is relatively more important than additional investments in period 1 and 3 of childhood, while they are more balanced in period 2. The production function is concave as  $\rho < 1$  and we find clear evidence of complementarities of investments across stages of childhood. The estimated elasticity of substitution ( $1/(1 - \phi) = 1.7$ ) implies that the technology of skill formation is strongly complementary in investments across periods. We also find evidence of dynamic complementarity at the optimum as cross derivatives  $\frac{\delta^2 h_t}{\delta x_i \delta x_j}$  are always larger than zero for tall combinations of periods  $i$  and  $j$ .

Parameter	Estimates
$\gamma_1$	0.5991
$\gamma_2$	0.4602
$\gamma_3$	0.6194
$\phi$	0.4282
$\rho$	0.7487
$\delta$	2.23

Table 3.2: Estimated Parameters for the Technology of Skill Formation

For robustness checks we estimate the technology parameters under different combinations of the main calibrated parameters. We try with different CRRA coefficients  $\sigma = \{0.5, 1.0, 1.5\}$ , we change the variance of income shocks using  $\sigma_\varepsilon^2 = \{0.1, 0.3, 0.8\}$  and we change the parameters of the Terminal Value Function  $(\eta, \tau)$ . Even though our estimates slightly change, the same patterns remain. Complementarity of investments across different periods of the childhood are persistent in the technology of skill formation.

We construct the analogous two-dimensional graphs we use for our data but with our simulations at the optimal structural parameters. In all the cases, our simple model is able to replicate the main finding that schooling is maximized when the timing of income is balanced. Figure 3.24 replicates the simulated schooling attainment if income is shifted from  $I_1$  to  $I_2$  holding  $I_3$  and  $PI$  at median values. Figures 3.25 shows the results when income is shifted from  $I_1$  to  $I_3$ . The full set of simulations are shown in panels B15-B16. In all cases simulated patterns of schooling are inverse U-shaped, and our tests reject monotonicity and horizontal slopes.

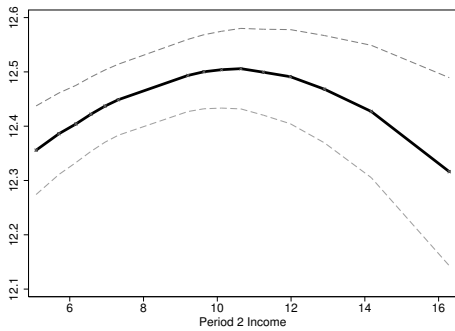


Figure 3.24: Simulated schooling attainment and Income age 6-11,  $PI$  and Income age 12-17 fixed at median values

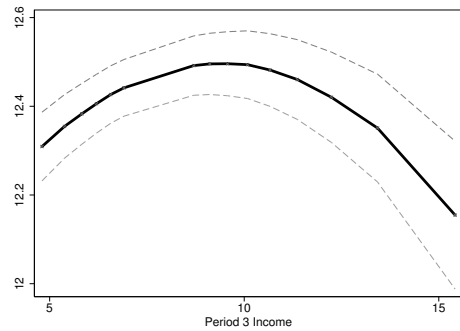


Figure 3.25: Simulated schooling attainment and Income age 12-17,  $PI$  and Income age 6-11 fixed at median values

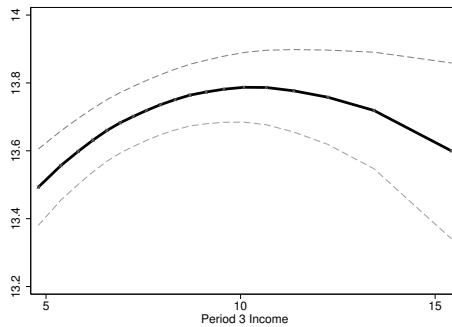


Figure 3.26: Simulated schooling attainment and Income age 12-17,  $PI$  and Income age 0-5 fixed at median values

Note: Simulated schooling attainment is shown with 95% confidence intervals. Simulated incomes in £ 10,000s. Plots based on 100,000 simulated income draws evaluated at the estimated technology of skill formation.

We use our simulations to investigate the mechanisms behind the inverse U-shapes. First we tackle the role of credit constraints. If parents are credit constrained, optimal investments should be reactive to income shocks, anytime the

technology is complementary in investments. In the next figures we study the reaction of investments to income shocks. Figure 3.27 shows the effects of shifting income from  $I_2$  to  $I_1$  holding  $I_3$  and  $PI$  fixed at median values. As parents are credit constrained, previous to the shift they were underinvesting in early childhood ( $x_1$ ) and over-investing in middle childhood ( $x_2$ ). As income shifts from middle to early childhood parents adjust their choices increasing investments in the first period and decreasing investments in the second period, while investments in adolescence remain non-reactive. Figure 3.28 shows the effects of shifting income from  $I_3$  to  $I_1$  holding  $I_2$  and  $PI$  fixed at median values. Parents take advantage of an increase in disposable income in early childhood borrowed from adolescent years and would invest more in early childhood and less in adolescence, while investments in middle childhood remain fairly flat. Figure 3.29 shows the effects of shifting income from  $I_3$  to  $I_2$  holding  $I_1$  and  $PI$  fixed at median values. Here investments in middle childhood increase, investments in adolescence decrease and investments in early childhood do not react as a consequence of the shift. The patterns are the same when incomes left aside are fixed at different deciles, as shown in Panels B17 and B18 in Appendix B.

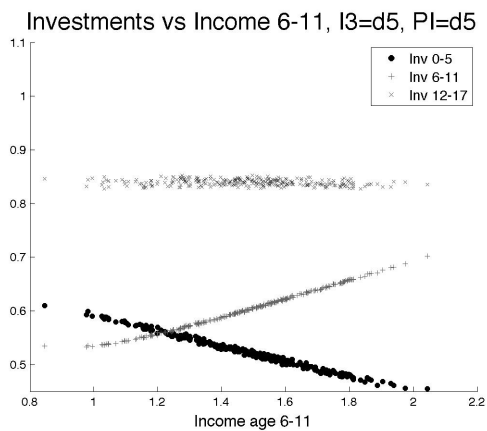


Figure 3.27: Simulated Investments and Income age 6-11,  $PI$  and Income 12-17 fixed at median values

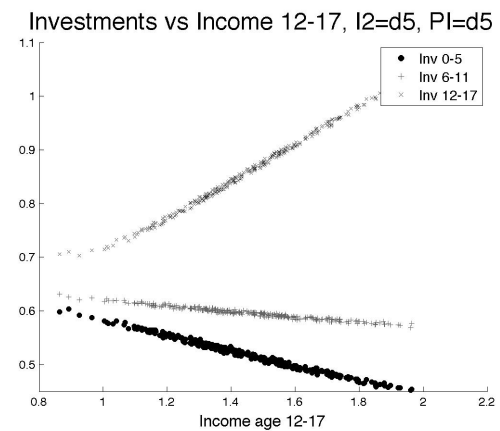


Figure 3.28: Simulated Investments and Income age 12-17,  $PI$  and Income 6-11 fixed at median values

We further investigate the role of uncertainty. Uncertainty in income shocks and a technology of skill formation which is complementary in investments may potentially produce that parents would not invest optimally in early childhood even if they are given all the money upfront. In other words, parents may over-save because of precautionary motives. This is supported by our findings about how

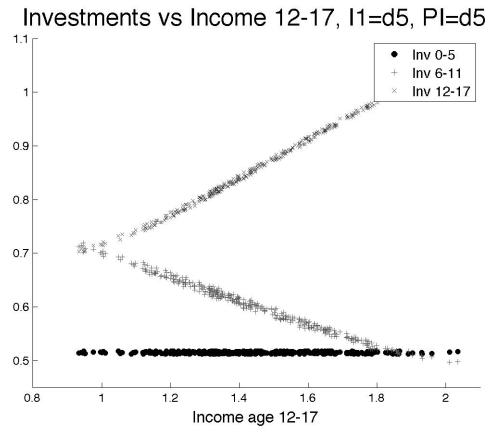


Figure 3.29: Simulated Investments and Income age 12-17,  $PI$  and Income 0-5 fixed at median values

Note: Simulated per-period investments. Simulated incomes in £ 10,000s. Plots based on 100,000 simulated income draws evaluated at the estimated technology of skill formation.

savings react to income shocks. Figure 3.30 shows how savings react to shifts in income from  $I_2$  to  $I_1$  holding  $I_3$  and  $PI$  fixed at median values. Savings in early childhood increase in reaction to income shocks, while savings in the future remain unchanged. Figure 3.31 shows that if income is shifted from  $I_3$  to  $I_1$  holding  $I_2$  and  $PI$  fixed at median values, then parents increase savings in both early and middle childhood for precautionary motives. Finally, Figure 3.32 shows that if income is shifted from  $I_3$  to  $I_2$  holding  $I_1$  and  $PI$  fixed at median values, parents over-save in middle childhood, while investments in early childhood are non-reactive.

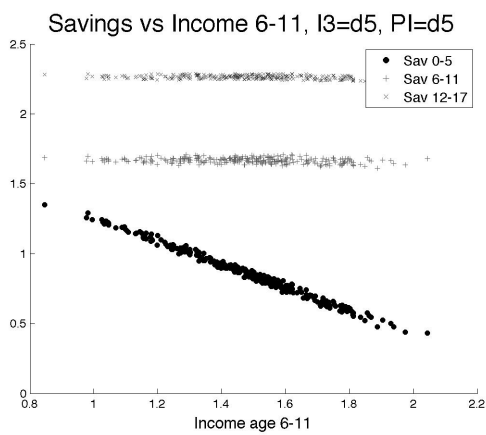


Figure 3.30: Simulated Savings and Income age 6-11,  $PI$  and Income 12-17 fixed at median values

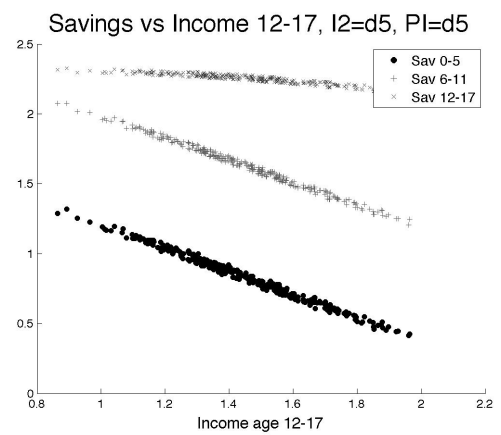


Figure 3.31: Simulated Savings and Income age 6-11,  $PI$  and Income 12-17 fixed at median values

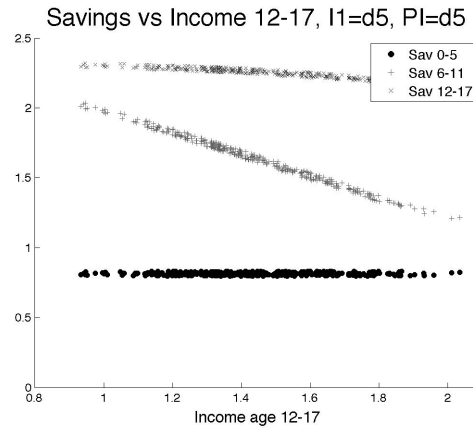


Figure 3.32: Simulated Savings and Income age 6-11,  $PI$  and Income 12-17 fixed at median values

Note: Simulated per-period savings. Simulated incomes in £ 10,000s. Plots based on 100,000 simulated income draws evaluated at the estimated technology of skill formation.

In summary when the technology of skill formation is complementary in investments, there is income uncertainty and parents are credit constrained, the timing of income is easily transmitted into human capital through investment decisions. The simulated effects of the timing of income in per-period investments suggest that a policy that alleviate credit constraints or facilitate inter temporal income transfers would prevent sub-optimal investments. This could partially explains the inverse U-shape of human capital outcomes when income is shifted across periods, particularly the downward sloping part of the data patterns for schooling outcomes. However, the role of uncertainty also plays a role beyond the existence of credit constraints. Income uncertainty leads parents to over-save in reaction to positive income shocks, which could harm their ability to invest in the human capital of their children optimally. This is a particularly useful argument to explain why we observe an upward sloping portion of the behavior of human capital when parents are able to borrow money from the future.

### 3.6.4 Sensitivity

The analysis performed before is of course conditional to a technology that is complementary in parental investments. Many of the preference parameters used in the model are not possible to be identified from our data, so we calibrate them. However, we test that the estimated technology is robust to the calibration by re-estimating the model under different scenarios of preferences, and lifting credit constraints.

Table 3.3 presents the estimated technologies for the different scenarios. In all cases the estimated elasticities of substitution of investments support a technology of skill formation which is complementary in investments.

Parameter	Benchmark	No Credit Con- straints	Low Elasticity of Consumption ( $\sigma = 0.95$ )	High Elasticity of Consumption ( $\sigma = 0.1$ )	High valuation of child's human capital relative to assets ( $\frac{\eta}{\tau} = 5$ )	Low valuation of child's human capital ( $\frac{\eta}{\tau} = 0.2$ )
$\gamma_1$	0.59	0.39	0.54	0.37	0.48	0.42
$\gamma_2$	0.46	0.69	0.48	0.53	0.63	0.53
$\gamma_3$	0.62	0.56	0.63	0.70	0.73	0.43
$\phi$	0.43	-0.31	-0.36	0.30	0.49	0.50
$\rho$	0.75	0.52	0.58	0.71	0.69	0.49
$\delta$	2.23	2.97	3.60	2.45	3.03	2.56
$1/(1 - \phi)$	1.75	0.76	0.73	1.42	2	2

Table 3.3: Sensitivity of the Estimated Technology of Skill formation under different calibrated preference parameters

### 3.7 Concluding Remarks

The question of whether early or late income is most productive in producing child human capital is important in order to further our understanding of how the stock of adult human capital accumulates. Should policy be targeted towards very young children, or adolescents, in order to reduce child inequalities in achievement later in life? Our dataset consists of the population of children born in Norway in the 1970s. The large scale of this allows us to estimate the relationship between child's education and the income received in their early years from 0-5, the middle period aged 6-11 and during adolescence aged 12-17, in a semi-parametric setting. The benefits of this choice of methodology are evident from the subtleties of the interactions between income in the three periods that can be evaluated. We find that early years income (age 0-5) is as important in general as income aged 6-11 and as income during adolescence. One exception to this rule is that for the outcome college attendance, an increase in adolescent income at the expense of early income will raise participation at college. There are complementarities across adjacent periods, suggesting that an even bundle of income across periods is optimal, compared to extreme levels. Put another way, when income in any period falls below a threshold, it does so to the detriment of child human capital. These results are

robust to the choice of bandwidth and to a semi parametric specification where we control for a detailed set of family inputs.

That we do not find a strategically different pattern for the poorer families, likely to face credit constraints is not entirely surprising. In their test of a prediction of the permanent income hypothesis, that contemporary consumption responds only to unpredictable changes in income, Shapiro and Slemrod (1995) and Souleles (1999) all find the same pattern for credit constrained households as all others.

Simulations of behavioral models of parental investments in children emphasizing credit constraints and income uncertainty shed light about why the timing of income matters for human capital formation. In the presence of credit market imperfections and income uncertainty, income shocks are transmitted into parental decisions about investments in human capital and savings. This is true when investments across different stages of childhood are complement, meaning that if parents do not invest optimally in early childhood because of those market imperfections, this has detrimental effects in the accumulation of skills that cannot be remediated later at the same cost.

There are policy implications from the results of our study. In fact, we find that even in the setup of a developed welfare state as in Norway in the 70's there is ground for public policies that provide insurance against income shocks and allow parents to borrow from their future earnings or even from their children's future earnings. And if these types of policies can alleviate the negative effects of income fluctuations in the long-term human capital formation in Norway, they may be even play a more important role in poorer countries.

# Chapter 4

## The Role of Beliefs in Parental Investments and Child Development

### 4.1 Introduction

There is abundant evidence from the literature on child development that the early years of an individual's life greatly influence one's life trajectory. Socioeconomic gaps in early childhood widen over the life-cycle in both developed as well as in developing countries (See Fernald, Weber, Galasso, and Ratsifandrihamanana (2011), Moon (2010) and Schady, Behrman, Araujo, Azuero, Bernal, Bravo, Lopez-Boo, Macours, Marshall, Paxson, and Others (2014)) so there is high potential returns to interventions supporting poor families and their investments in children. High quality early childhood programs have been shown to lead to gains in child development and even longer term outcomes, especially for disadvantaged children.

A traditional approach to deal with early childhood poverty (and its consequences) has been the provision of financial resources. The literature has recognized that this approach is largely insufficient (Paxson and Schady (2007), Heckman (2011), Cunha and Heckman (2009), Cunha, Elo, and Culhane (2013)). Even though poor parents do suffer from a lack of resources, when these are made available their use would greatly depend on parental beliefs and preferences towards child-rearing, alternative uses of resources (both time and money), information on the best use of resources, and expectations of returns to investments in children. Therefore, there are potential high returns to the improvement of parenting skills in poor families.

There is wide agreement among economists, child development specialists, policy makers, and international financial organizations, that more research is necessary to understand by how much and by which channels parenting affects the process of



human capital formation both in cognitive and non-cognitive dimensions.

Putting this in a broader context, there are two important challenges driving the current research agenda in early child development. First, it has become a priority to evaluate potentially scalable parenting interventions which are sustainable in the long-run by either involving the local communities in the implementation, or by nesting them in the formal structure of public services. Second, we need a deeper understanding of the factors, mediators and constraints by which better parenting translates into investment in children's human capital and child development. One way to do so is to evaluate experimentally multiple policies with many treatment arms in different contexts, each covering a different potential mediator or constraint, but this is costly and hard to implement at scale. A complementary approach to understanding child investments is to evaluate simpler and ongoing interventions which attempt to modify parental behavior, and use the experimental design together with very rich data to investigate those channels with an insightful empirical approach or by estimating realistic behavioral models of parental decisions in child investments; realistic in the sense that parents face multiple sources of uncertainty, multiple forms of investments, and child outcomes have multiple dimensions.

The objective of this work-in-progress research is to go beyond the analysis of how financial constraints impact parental investments in children, and to understand how parenting, in particular the role of parental beliefs and attitudes towards child-rearing, determines investments in children and their subsequent development. To do so, we rely on the impact evaluation of "Nadie es Perfecto" (NEP hereafter), a nationally-scaled group-based intervention nested in the Chilean public health system that provides parents from poor economic backgrounds with information and support about the benefits of positive parenting strategies, encourages improvements in home environments, and fosters parental self-esteem in the child-rearing task. Our main goal is to use the experimental data from this intervention to uncover the pathways and mechanisms relating parental beliefs and expectations are related to child investments, parenting practices, and child outcomes, and use this knowledge to evaluate cost-effective and potentially transferable ECD policies in the context of poor and middle income countries.

Our research innovates in four aspects, concerning both data and methods. First, we collect new data measuring parental beliefs and expectations, which are rarely available, along with standard measurements of parental investments and child outcomes. We consider three types of beliefs, which complement each other, and together form a much more complete core set of beliefs than ever considered before.

There are beliefs about how adequate are different styles of interacting with children, such as being authoritarian or being permissive. There are beliefs about one's own ability to be an effective parent, or perceived self-efficacy. And there are beliefs about the expected returns to cognitive and non-cognitive stimulation of children building on Cunha, Elo, and Culhane (2013), who develop a new instrument for the elicitation of parental perceptions regarding the returns to cognitive stimulation among low income mothers in the US. They recover maternal perceptions regarding cognitive stimulation by asking mothers for cognitive developmental milestones children should achieve under different scenarios for home stimulation. We adapt a new version of this instrument, previously piloted for the Chilean context, and extend it to two dimensions of parental stimulation: cognitive (language) and socio-emotional (nurturing and disciplinary strategies).

Second, collecting this new data in a setting where we observe exogenous changes in parental beliefs, allows us to identify their impact on different dimensions of parental investments and child development. Even with very detailed measurements of beliefs, there may be unobservable determinants of parental beliefs that affect both parents in children. Furthermore, we control for further sources of bias that are not usually accounted for like measurement error. To do so, we formulate a technology of skill formation of different dimensions of child outcomes, investments and beliefs, and we estimate it with a dynamic factor model as in Cunha, Heckman, and Schennach (2010). Differently from them, in our setup we do not need to impose some exclusion restrictions when we estimate the latent factors for investments and beliefs, because we can exploit the exogenous variation from the RCT.

Third, we use our factor model to investigate the links between parental beliefs and parental investments, and the links between investments and child outcomes, by using econometric mediation analysis proposed by Heckman and Pinto (2013). In this approach, we show that we can decompose treatment effects between indirect changes in child outcomes through changes in investments, and direct effects through changes in the productivity of those investments, under relatively weak assumptions given the extensive number and quality of measurements we have available. This way, we can minimize the set of unobserved inputs that can potentially make this decomposition difficult.

Fourth, we go beyond the understanding of the pathways in the technology and we formulate and estimate a dynamic structural model of parental investments in children to separately identify the role of preferences and expectations in shaping behavior (Cunha and Heckman (2007)). The estimation of such policy-invariant

structural parameters will ultimately be useful to assess the cost-effectiveness of the current policy and simulate alternative new policies that are more cost-effective. In our model, parents won't fully know the technology of skill formation but they will learn about it in several ways (through information provision, or just through the act of parenting itself, in a learning by doing framework as in Badev and Cunha (2012)). Here, the elicited perceptions with the instrument adapted from Cunha, Elo, and Culhane (2013) will allow us to separately identify the role in outcomes of maternal perceptions regarding the returns to investments from actual behavior.

Such a complete treatment of the determinants of parental investments in children is not available elsewhere in the literature, because it is extremely difficult to put together the elements just described. We are in a unique position to do so. This work will contribute to a unique understanding of the mechanisms through which parenting influences the lives of children, which is essential if we want to design new interventions in this field. The evidence systematically organized by Engle, Black, Behrman, de Mello, Gertler, Kapiriri, Martorell, and Young (2007) and Engle, Fernald, Alderman, Behrman, O'Gara, Yousafzai, de Mello, Hidrobo, Ulkuer, Ertem, and Others (2011) in the LANCET (2007, 2011), is weak on the effectiveness and cost-effectiveness of scaled-up parenting interventions and on the pathways through which they potentially work. This research focuses its contribution in addressing these two research gaps.

## 4.2 The intervention

Nadie es Perfecto is a group parenting intervention adapted by the Chilean government from Nobody's Perfect, a parenting education program developed by the Public Health Agency of Canada since the early 1980s. NEP has been selected as a cornerstone parenting intervention within their new nationwide early childhood policy, Chile Crece Contigo. The program has been gradually rolled it out within the country since 2010. NEP attempts to promote positive parenting skills in caregivers, the use of non-violent disciplinary strategies, helping caregivers to acquire new techniques oriented to manage child behavior and to foster positive parent-child relationships. The methodology aims at improving improve parental self-esteem and problem solving skills, as well foster social support and networks.

NEP targets mainly parents with children aged 0 to 5 who live in poor and isolated conditions. It takes place in a group setting of 6-8 weekly group sessions of 6-10 parents. The key innovation of the approach lies in a flexible curriculum that

can be tailored to the group interests and needs, combined with facilitators trained during 6-8 weeks in each of the NEP topics and in how to manage group dynamics. The premise of the intervention is that to translate knowledge and beliefs into real behavioral change, participants need not only to increase their knowledge about the optimal practices and about themselves as parents, but also to emotionally connect to the way themes are discussed with other parents facing similar problems.

An existing combined qualitative and quantitative evaluation (Chislett and Kennett (2007); Skrypnek and Charchun (2009)) of the Canadian NEP program shows a decrease in negative or punitive practices, and improved parental ability to cope with parenting stressors, problem solving ability and perceptions of social support. Yet, the evaluation fails to rigorously examine whether these changes in parenting practices translate into gains in child development outcomes. Furthermore, it also fails to account for the relevant role of parental beliefs in triggering changes in parental behavior.

NEP has several potential features that make it an attractive model to be evaluated, both in the Chilean context as well as for the early childhood development literature in general:

1. Scaled-up intervention: As highlighted in many of the LANCET (2007, 2011) meta-analysis on early childhood developments, there is a dearth of research of large scale programs that have the potential for having an impact of child development. NEP is rolled-out at the local health centers and the sessions are delivered by trained health workers. NEP uses the infrastructure and human resources already existing in the health network with no further monetary and organization costs beyond training and material printing.

2. Cost effectiveness: the standard version of NEP costs only 10% per family attended compared to home visits, while according to LANCET 2011, expected effect sizes lie between 25%-50% of home visits. Hence, NEP could imply a relative gain in cost-effectiveness of about 2.5-5 times. This is all based on the premise that parenting programs have significant impacts on child development, which is what we propose to test. Given the mixed evidence in the literature on parenting programs only providing information to parents, compared to more effective interventions that combine parent-child interaction, we developed an intensive version of the intervention as a new arm as part of the evaluation. This version of NEP (called NEP Intensive) adds sessions with direct interactions between parent and child focused on the importance of language and play. The intensive version could represent a gain in relative cost-effectiveness of about 4 to 9 times compared to home

	Cost/(Family*hour)	Expected Impact in Child development
<b>NEP</b>	<b>\$ 0.65</b>	<b>0.25 SD</b>
<b>NEP Intensive (+2 sessions, 33%)</b>	<b>\$ 0.87</b>	<b>0.5 SD (LANCET)</b>
<b>Home Visits</b>	<b>\$ 7.48</b>	<b>0.5–1 SD (LANCET)</b>

Figure 4.1: Cost-effectiveness *Nadie es Perfecto* vs Home Visits

visits as it costs only 1/3 more than the standard version (NEP basic) per family attended, implying an additional gain of about 3 times compared to NEP Basic (see Figure 4.1).

The meta analysis in the Lancet Review showed that parenting programs including the opportunity to practice with their children tend to show larger effect sizes (median of 0.46, range 0.04-0.97) than for those providing only information to caregivers (median 0.12, range 0.03-0.34). Even though the review shows more significant effects in home visits<sup>1</sup>, there is also positive evidence from community-based trials both including parent-child interactions and focused only on the caregiver, although stronger effects are found in the former case<sup>2</sup>. Finally, the evidence on scaled-up parenting only programs is still rather scarce and mixed<sup>3</sup>.

<sup>1</sup>Cooper et al (2009) reports significantly positive effects in South Africa measuring the quality of mother-infant interaction and infant attachment (range 0.24-0.86). Janssens and Rosemberg (2011) report significant improvement in child cognitive development in St. Lucia. Bentley et al (2010) report that home visits in India aiming to teach parents about complementary food, responsive feeding and play, improve significantly child mental development.

<sup>2</sup>Aboud and Akhter (2011) compare child outcomes between an intervention giving only information to parents on health and nutrition of children, to one including also coached practice with children. Results show a larger language development in the latter case, but in both cases maternal knowledge is significantly improved. Among interventions focused only on parents, Kagitcibasi (2009) reports that training mothers in Turkey improved children long-term outcomes as college attendance, educational attainment and higher status occupations. Al-Hassan and Lansford (2010) report that center-based parenting groups in Jordan significantly improved parental reports of cognitive and social activities with their children, along with discipline and knowledge.

<sup>3</sup>Engle et al (2011) report results from the Ecuadorian program “Educa a tu hijo” (Educate your child), which combines health care with a structured parenting program coordinated by the health sector and delivered at the community level. Children in the program had higher cognitive scores than those not in the program. Nodira et al (2009) report results from the Family empowerment program, a large-scale community-based health and nutrition program in Uzbekistan, which shows some significant improvements in parent skills and knowledge, but not effects in child development outcomes. Engle et al (2010) also report results from the parenting program Care for Development Intervention in Tajikistan and Kyrgyzstan, where children from treatment areas showed significantly higher communication, gross motor and social development.

## 4.3 Literature Review

This research builds up on the developmental and economic literature of the role of beliefs shaping parenting investments, on the policy perspective for the importance of investing in parenting interventions, and on the literature on structural behavioral models for policy evaluation.

### 4.3.1 Beliefs and attitudes about child rearing

NEP looks to encourage parents to improve cognitive and emotional stimulation of their children by changing parental beliefs and attitudes towards child-rearing, and by improving their mental health and perceived social support. Through these mechanisms, the program is expected to affect long-term outcomes like the home environment, parent-child interaction, and child development. The child development literature has highlighted two sets of beliefs and expectations: parental beliefs about the best ways to raise children, or the core set of beliefs about the role of parenting, and beliefs of how able parents think they are for child-rearing.

#### (i) core set of beliefs about the role of parenting and parenting styles.

Conditional on the belief that parents can activate actions on parenting, the influential work by Baumrind (1966, 1968, 1973, 1989), also also reviewed by Bornstein (2001), suggest that parents have ideas about raising children that translate into one of three parenting styles: authoritarian, permissive and authoritative. Authoritarian parents exhibit high levels of structure and control in children, many times involving harsh disciplinary strategies, but low levels of warmth and communication. The opposite is thought for permissive parents. Authoritative or democratic parenting combines high and balanced levels of warmth and control and it has been associated with positive child development outcomes. She found that children in authoritative homes performed better in standardized tests, were more friendly with peers, more independent and assertive, cooperative with parents, and more achievement oriented. In contrast, children from authoritarian homes were more hostile and/or shy with peers, more dependent on parents, and less achievement-oriented. Children in permissive homes exhibited very similar patterns as those from authoritarian homes. Baldwin (1948) finds that children from homes with high levels of control and democracy were more cooperative, engaging and curious.

In the context of our intervention, participating parents are challenged to revisit their parenting styles and mindset about parenting along many dimensions.

Given the promotion of positive parenting is a key objective of the intervention, we expect the NEP to induce program participants to move towards more democratic/authoritative parenting. However, in order to adjust to cultural contexts, we use the baseline survey to adjust the original scales to more suitable measures of the two underlying dimensions of parenting styles.

**(ii) beliefs about themselves as parents.**

This dimension of beliefs is grounded in social cognitive theory, aligning with Bandura (1995)'s theory, in which self-efficacy is "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations". In other words, self-efficacy is a person's belief in his or her ability to succeed in a particular situation. Parents with the same beliefs regarding their children might have different views about their own ability to translate beliefs into practices. Bornstein, Haynes, Azuma, Galperín, Maital, Ogino, Painter, Pascual, Pêcheux, Rahn, and Others (1998) show that high levels of perceived self-efficacy serve mothers to engage more intensely in the parent role investing more effort in parenting.

In our intervention, at least one third of the sessions target at promoting participants self-care and at improving their self-image as parents. In the group dynamics, parents learn to recognize and reinforce their strengths as individuals, how to avoid harsh self-judgement, and discuss some basic strategies that will help them solve daily problems and reduce parental stress. Another important channel by which NEP improves parental self-esteem and perceived self-efficacy is by creating networks of social support. Usually caregivers participating in the sessions come from very disadvantaged backgrounds with little or none social support. Participants usually create among themselves a supporting network with other parents that face similar problems, that usually remain active after the end of the intervention.

From our baseline data, we find strong associations between SES and beliefs, in line with the literature. Our hypothesis is then to understand whether parental beliefs are at the core of SES gradients in investments and child outcomes widely reported. In fact, our baseline data provides useful insights to this regard. For example, we find that higher SES caregivers tend to be more authoritative, or they tend to combine in a more balanced way warmth and structure ideas about parenting. Furthermore, they feel more self-competent in the child-rearing task and perceive a higher social support. In contrast, lower SES caregivers tend to be more authoritarian or permissive, perceive themselves less competent as parents, perceive

dramatically lower levels of social support and face higher parental stress<sup>4</sup>.

### iii) Perceived Returns to Investments

We study how subjective expectations about the returns to investments shape behavior. By sharing the personal experiences about parenting within a group setting, NEP is likely to raise parental awareness that their actions can affect specific domains of child development. This change in beliefs is an important pathway for sustained change over time, as the awareness of how one's own parenting affects your child (and the possibility of change) is likely to be internalized with the families even after the short exposure to the intervention.

If parental awareness is shifted as a consequence of the policy, we should observe a shift in the distribution of perceived returns to investments when comparing treatment and control groups. Subjective returns to investments will of course depend on whether parents believe intelligence is malleable or fixed, or the belief they can or cannot affect child academic achievement; and whether they believe temperament is malleable or not, or the belief they can affect emotional development. Wentzel (1998) shows that parents who believed that intelligence is malleable also believed they could affect their children academic achievement and had higher educational aspirations for their children. Teti, O'Connell, and Reiner (1996) show that mothers who believe they can or cannot change infant personality or intelligence modify their behavior accordingly. Parents may have a perspective of how fixed vs. malleable intelligence is: the cognitive dimension; or how fixed their personality/temperament is: the socio-emotional dimension.

Manski (2004) argues that the only way to separately identify preferences and subjective beliefs without assuming that people have specific expectations is to directly ask individuals their subjective probability distributions about outcomes. He also argues that subjective expectations should be retrieved in a probabilistic form because: i) probabilities have well-defined scales for responses and are interpersonally comparable, ii) we can use algebra of probabilities to examine the internal consistency of expectations, and iii) one can compare subjective probabilities with known event frequencies and check for the accuracy of those expectations.

Unfortunately, our psychology-grounded measure on parenting styles and perceived self-efficacy, while intended to capture awareness at least partially, are not a good starting point to recover subjective probability distributions. These self-reported scales, while showing high levels of internal consistency, were not built with the purpose of recovering probability distributions so it's hard to use them to

---

<sup>4</sup>See Annex 6.b: Baseline results.



infer expected returns to parental behavior. Because of this we build on an innovative instrument to elicit subjective beliefs about the returns to investments, which was specially built to recover probability distributions in the context of models of parental investments in children like in Cunha, Heckman, and Schennach (2010). The purpose of the instrument, adapted from Cunha, Elo, and Culhane (2013), is to resemble for the case of parental investments in children what has been done in the context of subjective returns to schooling since the 90's (Jovanovic (1979); Arcidiacono, Hotz, and Kang (2012), and Attanasio and Kaufmann (2009)). The instrument is explained in section 4.5.4

## 4.4 The Evaluation Design

### 4.4.1 Target Population

The target population of clinics includes basic family health care, rural health clinics, urban health clinics and health establishments with minimal service complexity: this corresponds to about 600 clinics in Chile belonging to the primary health care network, covering 342 municipalities. While Chile is a middle-upper income country, the intervention draws from the poorest section of the population, with more than half of the target population in the bottom quintile of the income distribution. The primary health care system in Chile is accessed only by the bottom income quintiles. In fact, more than 50% of the target beneficiaries live with less than US\$ 90 a month per capita. Moreover, 30% of the children in our sample exhibit socio-emotional problems and language delays in the baseline survey. Most of these children are concentrated among families belonging to the poorest 50% of our sample. Therefore, the magnitude of the developmental risks among the poorest household in Chile, which are the ones targeted by NEP, makes the analysis relevant from a developmental perspective and salient for other middle income countries. For more details, see Annex 10.1.

### 4.4.2 Sampling

A three-stage clustered sampling strategy was implemented. In the first stage, a representative sample of clinics stratified by the type of health center was chosen from a set of 612 health centers located in both urban and rural areas all over the country. The stratification was conducted fixing the fraction of rural and urban clinics, and the type of clinics which included family health centers, general health centers and

small hospitals, which may differ in the infrastructure and human resources available to deliver the program.

In the second stage, within each health center a sample of 18 families was randomly drawn from a potential wait-list of participants (usually between 45 and 60 potential participants per center). These long wait-lists were formed by the facilitators previous to the baseline survey. Families were put on waiting lists after a regular health visit to the center, conditional on satisfying the inclusion criteria for eligibility into the program as assessed by the health professional.

Finally, in a third stage, the 18 families selected to participate in the evaluation were randomly allocated to treatment and control groups. 6 families were offered NEP basic, other 6 families were offered to participate in NEP intensive, and a third randomly selected set of families served as a control group. They were offered to participate in NEP but only after the finalization of the evaluation, and they signed and consent form authorizing this delay. The process of random assignment to treatment groups was web-based and centralized in the Ministry of Health, so no one apart from our fieldwork coordinator had control over the randomization process. The compliance to the assignment within centers was strictly monitored in conjunction with the local program administrators.

This is a very natural way to design the evaluation, taking advantage of the limited program coverage. It implies only a very small change in the way families are selected to participate, and it has very limited ethical problems. It gives us the possibility of having a very rigorous study, where treatment and control populations will be identical by virtue of randomization. It is an evaluation design that is a consequence of a perfect marriage of implementation and evaluation concerns right from the start of the program.

### 4.4.3 Power Calculations

Each parenting course is given to a group of 8 to 12 parents. It was agreed that each facilitator would invite about 20 people, and from these 20 a random sample of 6 will be used for the study. The expectation was that, at least three out of these six would effectively participate in the program. Similarly there will be about 20 parents in the control group for each facilitator, of which 6 parents were sampled for the study. Our original goal was to detect an impact of the program of at least 0.25 SD on different indicators of parental beliefs and practices, and child development  $Y_i$ .

We expected compliance rates of roughly 40% among individuals randomized

in and full-compliance among randomized out. This implied a total sample of 162 centers and 2,916 families evaluated. Due to implementation problems, a sample of 20 centers were not able to participate in the evaluation, and compliance rates were roughly 35%, which led us to decrease our expectations about the impacts to be detected for the follow-up survey. Under our new power calculations, we would be able to detect an impact of the program of 0.35 SD on different indicators of parental beliefs and practices, and child development, which is still in line with the findings of the literature.

According to The Lancet series, even though an effect size of 0.35SD is above the median for interventions only working with parents, it is far below the median for interventions including parent-child interactions, which is one of the main targets of our impact evaluation design. Therefore, we are confident that at least in NEP Intensive we will be able to easily detect effect sizes that are relevant from an economic point of view.

Nevertheless, we have adopted two strategies to improve the precision of our estimates. First, we include measures of maternal IQ and personality traits. These variables have been reported to be highly predictive of child development outcomes in the Chilean context. And second, an important reason why we collected such a detailed baseline survey is that we can improve even more the precision of our estimates by estimating value added models including the baseline test scores as controls. The inclusion of past test scores should absorb an important proportion of the variance of our treatment effects.

Adopting these two approaches, we are confident we will be able to detect at least our original effect sizes of 0.25 SD for our ITT parameters, or even lower

To incorporate real compliance rates into power calculations we follow an IV approach to the computation of standard errors. In this approach,  $T_i = \{0, 1\}$  is the a dummy variable taking value 1 if the caregiver effectively received treatment or not, and  $Z_i = \{0, 1\}$  is the original random allocation. We want to estimate the following model for the effect of the intervention

$$Y_i = \beta_0 + \beta_1 T_i + u_i$$

We use  $Z_i$  as an instrument for effective treatment. This is a good instrument as long as  $cov(Z, u) = 0$ , which is ensured by random allocation process, and  $cov(Z, T) \neq 0$ , which is satisfied as long as a big enough fraction of individuals randomized in do comply with treatment. In a standard clustered approach, where  $N_c$  is the number of clusters (health centers) and  $M$  the size of the cluster equal for control and treatment

Nc/M	1	2	3	4	5	6	7	8
25	2.24	1.57	1.28	1.10	0.98	0.90	0.83	0.77
50	1.57	1.10	0.90	0.77	0.69	0.63	0.58	0.54
75	1.28	0.90	0.73	0.63	0.56	0.51	0.47	0.43
100	1.10	0.77	0.63	0.54	0.48	0.43	0.40	0.37
125	0.98	0.69	0.56	0.48	0.43	0.39	0.36	0.33
150	0.90	0.63	0.51	0.43	0.39	0.35	0.32	0.30
175	0.83	0.58	0.47	0.40	0.36	0.32	0.30	0.28
200	0.77	0.54	0.43	0.37	0.33	0.30	0.28	0.26

Table 4.1: Power Calculations

arms, it can be shown that the variance of the IV estimator is

$$\text{var}(\beta_{IV}) = \frac{\sigma^2}{N_c M} * (1 + (M - 1)\rho_u) * \frac{1}{\text{corr}(Z, T)^2} \quad (4.1)$$

Where we consider an intracluster correlation  $\rho_u = 0$ , compliance rates of 35%, and the variance of outcomes,  $\sigma^2$ , has been standardized to 1. We compute the number of centers and the sample size per treatment arm in each center for a level of significance of  $\alpha = 5\%$  and a power of the test  $p = 80\%$ . The results are summarized in Table 4.1:

Under these assumptions, in using 6 individuals per facilitator and treatment arm we included 150 facilitators in order to detect an effect size of about 0.35 SD. This implies a sample size of 150 (facilitators) \* 6 (parents per facilitator and group) \* 3 (treatment 1 (basic), treatment 2 (intensive) and control groups) = 2,700 (parents). In sum, the sample was comprised of: 150 health clinics, stratified by type of health center, rural/urban; 300 facilitators (150 for the basic NEP and 150 for the enhanced NEP intensive); 18 households total per health center (6 treatment NEP basic + 6 NEP intensive + 6 control).

#### 4.4.4 Measurements

The impact evaluation has been designed as a randomized control trial with two observations over time: a baseline survey before the intervention, which took place during June-September 2011, and a follow-up survey 18 months after the intervention (Oct 2013-March 2014). The target population are mothers and caregivers of children 0-6 that are participating in the Chile Crece Contigo system. Most of the variables and instruments described below were already applied in the baseline survey, and we will explicitly mention those who are added to the follow-up to estimate the behavioral model.

A first group of variables measure different dimensions of parental beliefs, attitudes and expectations. The first two groups of scales are our psychology-grounded

measures of parental beliefs, while the third one refers to the adapted instrument to recover parental perceptions about returns to investments.

### Measures of beliefs and expectations:

1. Beliefs of parental self-efficacy: we have already collected data on the baseline survey with standard instruments measuring sense of competence (Parenting Sense of Competence Scale (Ohan, Leung, and Johnston (2000)) and parental stress (Parenting Stress Index, Short form).
2. Core set of beliefs about parenting: We measure ideas about structure and warmth and the three parenting styles proposed by using Ideas About Parenting (IAP) questionnaire. This measure can be used to characterize parenting in terms of authoritarian, authoritative, permissive, and neglectful. In order to adapt the scale to the Chilean cultural context we used Item Response Theory to translate the original parenting styles into the two dimensions of parenting proposed by Baumrind (1968). We construct two sub-scales, warmth and structure, which are the ones used in the model.
3. Instrument to Elicit subjective expectations about the returns to investments: we adapt the scale to elicit beliefs of Cunha, Elo, and Culhane (2013), which has been pre-tested in November 2012 to the sample of NEP with the target of eliciting parental beliefs regarding the benefits of providing children better cognitive stimulation, and the benefits of using non-violent/harsh disciplinary strategies to manage child behavior. (Further details in section 4.5.4).

### Measures of parental investment:

1. Non-cognitive stimulation: In the baseline we used two sub-scales of the Parent Behavior Checklist (Fox, 1994).
  - (a) Nurturing, Communication and Socio-emotional stimulation: the Parents were asked to indicate how frequently, rated along a 5 point scale ranging from “never” to “many times each day”, they engaged in 16 different activities with their child over the past couple of weeks.<sup>4</sup> Example items include: “How often did you and your child laugh together?”; “How often did you play games with your child?”; and “How often did you praise your child for learning new things?” Higher scores reflect more frequent engagement in nurturing parenting behavior.

- (b) **Discipline Practices:** Parents were asked to indicate how frequently, rated along a 5 point scale, ranging from “never” to “many times each day,” they behaved in a variety of ways when their child broke the rules or did things the parent did not like. Example items include: “Ignore it, do nothing”; “Spank your child”; “Use time out.” Higher scores indicate greater frequency of engaging in that type of response to children’s misbehavior.
2. **Cognitive stimulation:** In the baseline we used the Family Care Indicators (Unicef, validated by Hamadani, Tofail, Hilaly, Huda, Engle, and Grantham-McGregor (2010)), which measures the quality time spent with children in learning and playing activities for young children at home. Examples of questions are how often parents take children out to the park, cinema or other recreational activities, whether there is always an adult looking after children, the frequency of learning and play activities with children, and the amount and variety of play and learning materials. In the follow-up we will pre-test a revised version of the FCI and the HOME-SF (play material and cognitive stimulation components) to be adapted to the older age group, as well as an adapted HOME scale by Aboud (2006), that combines self-report and observation.

### **Measures of child development outcomes:**

We consider those developmental domains that are expected to be affected by the intervention, and which have been described in the literature to have predictive power on adult measures of human capital :

1. **Language:** We have measured both receptive and expressive language. We use the Spanish version of the Preschool Language Scale (PLS-4). The PLS-4 has been preferred to the commonly used PPVT (Peabody Picture Vocabulary Test) because it includes both receptive and expressive vocabulary and because it allows evaluating all age groups relevant for the intervention.
2. **Executive functions:** these are the cognitive aspects of self-regulation (Blair and Ursache (2011), Blair and Razza (2007)) and defined as working memory, inhibitory control, and attention shifting, which have been proved to be important predictors for children’s social and academic development. We will continue using the Dimensional Change Card Sort (DCCS) task, appropriate for longitudinal uses starting from age 2  $\frac{1}{2}$  until young adulthood. For

younger children, we use the A-not-B task (Diamond (1985)) which captures both working memory and inhibition.

3. Maladaptive Behavior: There is conclusive longitudinal evidence that interventions working with children with highly disruptive behavior had positive implications in future labor market outcomes, decrease in criminal rates and improvements in social skills. Qualitative evidence from focus group discussions and feedback from the NEP facilitators have highlighted a high demand of caregivers to treat children behavioral strategies in the sessions. Program beneficiaries are looking for tools for solving behavioral issues with their children so we include this non-cognitive measure as one of the outcomes. We measure maladaptive behavior by using the CLBC Achenbach Child Behavior Checklist, which captures internalizing and externalizing behavioral problems.
4. Socio-emotional development: In order to measure positive dimensions of how the child establish personal relationships with peers and with adults, and to encourage comparability with other nationally representative surveys on the same target population. The main instrument is the Battelle scale, social-personal dimension. It measures positive social development of children, as reported by their caregiver.

## Parental Traits

We include the following set of cognitive and non-cognitive parental traits:

1. Maternal depression: We collected data on symptoms of depression using the widely Center for Epidemiologic Studies Depression Scale (CESD).
2. Maternal Distress : Measured with the Parenting Stress Index
3. Perceived Social Support: Preliminary structured interviews in Chile from the group interaction lies in the perceived social support that beneficiaries report to have improved as a result of their participation to the group parenting program.
4. Maternal IQ : Maternal IQ has been proved to be highly correlated to child skills. We adopt the Wechsler Adult Intelligence Scale (WAIS) which has already been validated for the Chilean population in the nationally representative child development survey (Encuesta Longitudinal de la Primera Infancia, ELPI). The scale has two sub-dimensions: Language and Digit Span.

5. Personality Traits: Measured by the “Big Five” test, including the sub-scales of Extraversion, Agreeableness, Conscientiousness, and Openness and lower levels of Neuroticism. All these sub-scales have been found associated with cognitive and non-cognitive child skills.

### **Administrative and socioeconomic data**

Socio-demographic characteristics of the family are obtained from a socioeconomic survey. We collect rich data for all the household members about labor and non-labor incomes, transfers, family composition, employment status, wealth and housing conditions, access to health and community services, disability and health shocks. The baseline survey shows that NEP works with the poorest segments of the income per capita distribution, emphasizing the potential of the policy to alleviate early disadvantages.

The survey data will also be merged to detailed administrative data on each child-mother pair. Crece Contigo has a unique feature of a longitudinal follow-up of each child with key administrative data on maternal pregnancy, birth outcomes and health visits linked for each child. It will be therefore possible to access retrospective history on risk factors and health outcomes for each child by collecting information on the personal national identification number. An informed consent was administered to parents to be able access the health records of their children in the health clinics. The detail on the key retrospective outcomes will be obtained from the clinical folder for each child/mother dyad includes maternal psychological evaluations, postnatal depression scales, and child’s anthropometrics.

## **4.5 Treatment Effects**

In this section we describe how treatment effects are estimated and how we can decompose changes in child outcomes. We will assume a linear framework because it facilitates reliable estimation in small samples. The analysis is divided in four parts. First, we discuss the overall estimation of treatment effects accounting for selection bias. Individuals are selected at random either to participate in NEP or to go on a wait-list, therefore selected individuals can then decide or not to participate in the program. The offer of the program is randomly assigned in a population of eligible individuals, but the take-up of the program is not random.

Second, we discuss how potential changes in child outcomes due to treatment will be decomposed by using Mediation Analysis (Heckman and Pinto (2013)). In



this framework, parental beliefs that are directly affected by the policy can affect outcomes in three possible ways: i) an indirect effect by changing measured parental investments; ii) a direct effect changing unobserved parental investments; and iii) a direct effect changing the technology that maps investments into output, i.e., the productivity of such investments. We discuss under which assumptions we can separately identify these channels using our data.

Third, most of our outcomes and inputs are measured with error, potentially biasing the estimation of overall treatment effects and the decomposition. We propose a latent factor model from our set of measurements, and we discuss the conditions for identification of the latent factors. After identifying the latent factor model we discuss estimation.

Finally, we propose the estimation of a full behavioral model in which parental beliefs will have two central roles in explaining how parents invest in children. First, our psychology-grounded measures of beliefs about how to raise children and about themselves as parents will drive both the level of investments and the productivities of those investments. Second, parents will have subjective expectations about the technology and we collect data to elicit them. This will allow us to identify a model in which there is heterogeneity both in preferences and expectations, and where parents learn about the consequences of their actions.

#### 4.5.1 Identifying treatment effects

Let  $t$  denote the timing of the baseline data and  $t + 1$  the follow-up data.  $\theta_{d,t+1}^O$  is an error-free outcome of interest at follow-up which can vary with treatment arm  $d$ . In the context of our study, we call outcomes to any variable that can be affected by treatment. Therefore, they can include inputs like beliefs ( $B$ ) or investments ( $I$ ), or outputs like child skills ( $K$ ). Imai, Keele, Tingley, and Yamamoto (2011) present the idea of Sequential Ignorability, under which counterfactual outputs and counterfactual inputs are independent of treatment conditional on pre-program variables. For our purpose, we have three types of pre-treatment variables: ii)  $\theta_t^O$  are error-free measured outcomes at baseline; i)  $\theta_t^P$  are error-free maternal cognitive and personality traits; and iii)  $X_t$  is a vector of household characteristics. Finally, let  $D$  an indicator variable denoting participation in NEP, and  $Z$  is an indicator variable denoting whether an individual was invited to enroll in NEP, or whether he was put on a wait-list. The treatment effects on outcome  $O = \{K, I, B\}$  is recovered from the following two-stage estimation

$$\begin{aligned}\theta_{d,t+1}^O &= \tau_0 + \tau D + \sigma_d \theta_t^O + \alpha_d \theta_t^P + \beta_d X_t + \varepsilon_{d,t+1} \\ D &= a_0 + a_1 Z + a_2 \theta_t^O + a_3 \theta_t^P + a_4 X_t + \eta_t\end{aligned}\quad (4.2)$$

where we allow treatment to change directly the effect on outcomes, but also change the productivity of pre-treatment variables. In this setup, the gains from treatment are then represented by

$$E[\theta_{1,t+1}^O - \theta_{0,t+1}^O] = \tau + (\sigma_1 - \sigma_0)E[\theta_t^O] + (\alpha_1 - \alpha_0)E[\theta_t^P] + (\beta_1 - \beta_0)E[X_t] \quad (4.3)$$

Notice that the allocation to treatment  $Z$  needs to be a good predictor of program participation  $D$ , i.e., conditional on being invited, the take-up of the program is high. Low take-up rates may lead to large standard errors in the estimates of  $\tau$ ,  $\sigma_d$ ,  $\alpha_d$ , and  $\beta_d$ . Moreover, if these estimates vary in the population, and there is non-compliance only for those who are invited to participate in NEP, but who never take it up,  $E[\theta_{1,t+1}^O - \theta_{0,t+1}^O]$  measures the average treatment on the treated. An analysis of the take-up rates allow us to rule out non-compliance for those who are not invited to participate, but who make their way into NEP.

#### 4.5.2 Decomposing treatment effects

##### Changes in child outcomes

We want to go beyond the estimation of treatment effects and we attempt to investigate the mechanisms by which the policy can potentially impact child outcomes. As shown above, the treatment effects of the program on child outcomes can be estimated under weak assumptions. However, these estimates tell us little about which mechanisms may be important for producing impacts on children. Under slightly stronger assumptions it is possible to make substantial progress on the issue of mechanisms, even without a fully specified behavioral model.

We are interested in estimating the channels through which NEP affects child outcomes. Our hypothesis is that NEP changes exogenously parental beliefs and expectations, which in turn can influence child outcomes in three possible ways: i) indirectly encouraging measured stimulating parenting practices (change in investments); ii) directly changing child outcomes by affecting how productive those investments become; and iii) shifting other unobserved inputs. We investigate such mechanisms by estimating the following linear technology of skill formation

$$\theta_d^K = \tau_d + \tau_d^I \theta_d^I + \tau_d^B \theta_d^B + \sigma_d \theta^K + \alpha_d \theta^P + \beta_d X + \varepsilon_d \quad (4.4)$$

where for simplicity of notation we have dropped the time indexes. In this formulation,  $\theta_d^K$  are child cognitive and non-cognitive outcomes measured at follow-up  $K = \{C, N\}$ ;  $\theta_d^I$  are parental investments in children measured at follow-up, where for now we consider only one dimension of stimulation;  $\theta_d^B$  are parental beliefs measured at follow-up, where beliefs are our psychology-grounded measures of parenting styles and perceived self-efficacy, and the elicited maternal perception about the returns to investments; finally,  $\theta^K$  are child cognitive and non-cognitive skills measured at baseline. Under this specification, participation into treatment can shift measured inputs like beliefs and investments, can change unobserved inputs ( $\tau_d$ ), and can also induce a change in the productivity of inputs and pre-treatment variables.

In the standard potential framework, we can write  $\theta_d^K$  as

$$\theta_d^K = D\theta_1^K + (1 - D)\theta_0^K$$

and similarly

$$\begin{aligned} \theta_d^I &= D\theta_1^I + (1 - D)\theta_0^I \\ \theta_d^B &= D\theta_1^B + (1 - D)\theta_0^B \\ \varepsilon_d &= D\varepsilon_1 + (1 - D)\varepsilon_0 \end{aligned}$$

so the treatment gains are

$$\begin{aligned} E(\theta_1^K - \theta_0^K \mid \theta^K, \theta^P, X) &= (\tau_1 - \tau_0) \\ &+ \Delta\tau_d^I E(\theta_0^I) + (\tau_0^I + \Delta\tau_d^I) E(\theta_1^I - \theta_0^I) \\ &+ \Delta\tau_d^B E(\theta_0^B) + (\tau_0^B + \Delta\tau_d^B) E(\theta_1^B - \theta_0^B) \\ &+ (\sigma_1 - \sigma_0)\theta^K + (\alpha_1 - \alpha_0)\theta^P + (\beta_1 - \beta_0)X \end{aligned} \quad (4.5)$$

That is to say, the program can affect inputs (for example through a change in their expected returns to such inputs), and the productivity of inputs (for example, by teaching parents how to better use their time with the child). In the case of

inputs that can change with treatment (like investments and beliefs), we can actually decompose the gains into pure changes in quantities, pure changes in productivity, and the interaction between changes in quantities and productivities.

### Changes in inputs

In order to decompose the gains in the equation above, we need to understand what drives parental investments and beliefs, and to what extent the change in beliefs induced by the program can explain the change in parental investments in children. Let  $E(\theta_1^I - \theta_0^I)$  be the impact of treatment on parental investments and  $E(\theta_1^B - \theta_0^B)$  the impact of treatment on parental beliefs. In the same way we did it for child skills, we can write the policy functions for investments and beliefs as

$$\begin{aligned}\theta_d^B &= \tau_d^B + \sigma_d^B \theta^B + \alpha_d^B \theta^P + \beta_d^B Z + \omega_d^B \\ \theta_d^I &= \tau_d^I + \tau_d^I \theta_d^B + \sigma_d^I \theta^I + \alpha_d^I \theta^P + \beta_d^I Z + \omega_d^I\end{aligned}\quad (4.6)$$

where  $\theta^B$  and  $\theta^I$  are the measurements of beliefs and investments at baseline, and  $Z$  and  $X$  can have common variables.

We will assume that the marginal productivities of these equations are invariant to treatment (we can relax that assumption later). Then we can compute the treatment gains as

$$\begin{aligned}E(\theta_1^B - \theta_0^B) &= (\tau_1^B - \tau_0^B) + \sigma^B E(\theta_{bas}^B) + \alpha^B E(\theta^P) + \beta^B E(Z) \\ E(\theta_1^I - \theta_0^I) &= (\tau_1^I - \tau_0^I) + \tau^I E(\theta_1^B - \theta_0^B) + \sigma^I E(\theta_{bas}^I) + \alpha^I E(\theta^P) + \beta^I E(Z)\end{aligned}$$

which can then be used to estimate the decomposed treatment gains for child skills.

### Changes in productivities

As discussed above, NEP not only affect inputs but also has the potential to change the returns to investments and beliefs. In that case, equation 4.4 becomes analogous to the following linear technology

$$\theta_d^K = \tau_d(\theta_d^B, \theta^P, Z) + \tau_d^I(\theta_d^B, \theta^P, Z) \theta_d^I + \sigma_d(\theta_d^B, \theta^P, Z) \theta^K + \beta_d(\theta_d^B, \theta^P, Z) X_d + \varepsilon_d \quad (4.7)$$

where now the productivities to each of the technology inputs are a function of pre-program variables, and a function of parental beliefs.

$$\begin{aligned}
\tau_d(\theta_d^B, \theta^P, Z) &= \mu_d^\tau + \nu_d^\tau \theta^P + \eta_d^\tau \theta_d^B + \zeta_d^\tau Z + \rho_d^\tau \\
\tau_d^I(\theta_d^B, \theta^P, Z) &= \mu_d^I + \nu_d^I \theta^P + \eta_d^I \theta_d^B + \zeta_d^I Z + \rho_d^I \\
\sigma_d(\theta_d^B, \theta^P, Z) &= \mu_d^\sigma + \nu_d^\sigma \theta^P + \eta_d^\sigma \theta_d^B + \zeta_d^\sigma Z + \rho_d^\sigma \\
\beta_d(\theta_d^B, \theta^P, Z) &= \mu_d^X + \nu_d^X \theta^P + \eta_d^X \theta_d^B + \zeta_d^X Z + \rho_d^X
\end{aligned} \tag{4.8}$$

### Identification of the mechanisms

As noted by Heckman and Pinto (2013), estimating the equation (4.7) is problematic if observed and unobserved inputs are correlated. In that case, measured inputs are not statistically independent of counterfactual child outcomes conditional on treatment status and pre-treatment variables. In our case, this means that the effect of both investments and beliefs in child outcomes would measure a combination of the impact of these inputs, with a projection of them on correlated unobserved inputs  $\tau_d$ .

Identification of the effects of beliefs on investments, and the effects of beliefs and investments on outcomes becomes difficult in this case. This happens because of three reasons: 1) outcomes can change both because of changes in the returns (which are a function of beliefs), and because of direct changes in inputs; 2) observed and unobserved inputs are correlated, and unobserved inputs are also a function of changes in beliefs; and 3) unobserved inputs change with treatment too.

To achieve identification, one needs to minimize the number of unobserved inputs potentially correlated to observed inputs that vary with treatment. We follow two strategies. First, we collect a very rich set of inputs both at baseline and at follow-up that can vary with treatment. Regarding beliefs, we include psychology-grounded measures of parental beliefs about how to raise children, perceived self-efficacy, perceived social support, and we elicit perceived returns to investments, instrument which provides an additional powerful of identification of productivities (the instrument is described in section 4.5.4) In terms of parental investments, we collect extensive data about the home environments, the quality of time spent in nurturing activities with children, disciplinary strategies, expenditures in children, child care, and we merge our data with administrative records informative of the child's health care and the use of social services. Finally, we collect information about important

parental mediators like depression and parental stress. With such an amount of measures, potentially unobservable inputs that may be changing with treatment are very few. And second, even if there are some unobservable inputs which are left behind, we have incorporated baseline measures in all the production functions (for both inputs and outputs), so we can control at least for all unobservable inputs that time invariant. Given that the time window between the baseline and follow-up surveys is 18 months, these two strategies reduce the number of unobservable inputs varying with treatment to a reasonably low level so our assumptions, while stronger, are likely to hold.

### 4.5.3 The latent factor model

#### Measurements

Child outcomes, investments, beliefs and parental traits are likely to be measured with error. It is standard in the literature of cognitive and non-cognitive skills to use more than one measure to identify a single latent construct. This is done by specifying a latent factor model in which the signal of each measure is splitted from the error in order to identify the distribution of a latent factor. This methodology might involve some concerns though. One the one hand, if different measures supposedly identifying a single latent construct are highly correlated with each other, then measurement error is not a problem, as they are all noisy measures of the same outcome. On the other hand, if the correlation is too low, we are potentially forcing different measures to identify a single latent construct when in reality they are measuring different things. 4.2 indicates that correlations among the different child outcomes are not too high but also not too low, which justifies the assumption of at least one underlying latent factor. However, relatively high correlations between cognitive and non-cognitive measurements require to study correlations among subscales in order to make sure we can really identify two and not only one single latent factor for skills.

Child Outcomes	Language	Exec. Functions	Disruptive Behavior	Socio-emotional
Cog - Language	1			
Cog - Executive Functions	0.2999	1		
Non Cog - Disruptive Behavior	-0.1451	-0.1083	1	
Non Cog - Socio-emotional	0.4177	0.1653	-0.1217	1

Table 4.2: Correlations measures of cognitive and non-cognitive skills

Similarly, 4.3 and 4.4 show the correlations among measures of investments and beliefs, respectively. The magnitude of the correlations are not high enough to assume that these measurements are just noisy signals of the same latent construct, so their inclusion in a latent factor model is justified.

Parental Investments	HOME test	PBC Nurturing	PBC Neg.Discipline	PBC Pos.Discipline
HOME test	1			
PBC Nurturing Practices	0.4337	1		
PBC Harsh Discipline	-0.1354	-0.1498	1	
PBC Positive Discipline	0.1622	0.3447	0.185	1

Table 4.3: Correlations measures of parental investments and practices

Parental Beliefs	Autoritarian Style	Autoritative Style	Permissive Style	Perceived Self-efficacy
Autoritarian Style	1			
Autoritative Style	0.4395	1		
Permissive Style	0.237	0.3441	1	
Perceived Self-efficacy	0.0531	0.3052	0.0791	1

Table 4.4: Correlations measures of parental beliefs

## Identification

The first step before estimation is to summarize multiple measures of child development and parental investments into a low dimensional vector of factors and account for measurement error. Our measurement system is:

$$\begin{aligned}
 M_{d,j,t}^K &= v_{j,t}^K + \varphi_{j,t}^K \theta_{d,t}^K + \eta_{j,t}^K \\
 M_{d,j,t}^I &= v_{j,t}^I + \varphi_{j,t}^I \theta_{d,t}^I + \eta_{j,t}^I \\
 M_{d,j,t}^B &= v_{j,t}^B + \varphi_{j,t}^B \theta_{d,t}^B + \eta_{j,t}^B \\
 M_{d,j,t}^P &= v_{j,t}^P + \varphi_{j,t}^P \theta_{d,t}^P + \eta_{j,t}^P
 \end{aligned} \tag{4.9}$$

In this system,  $M_{d,j,t}^K$  denotes the  $j$ -th measurement of cognitive or non-cognitive child outcomes  $\theta_{d,t}^K$ ;  $M_{d,j,t}^I$  is the  $j$ -th measurement of cognitive or non-cognitive parental stimulation  $\theta_{d,t}^I$ ;  $M_{d,j,t}^B$  is the  $j$ -th psychology-grounded measurement of parental beliefs  $\theta_{d,t}^B$ ; and  $M_{d,j,t}^P$  is the  $j$ -th measurement of parental traits  $\theta_{d,t}^P$ . We should notice that  $\theta_{d,t}^K$ ,  $\theta_{d,t}^I$ ,  $\theta_{d,t}^B$  and  $\theta_{d,t}^P$  can be vectors or scalars, which means that we estimate flexible specifications where each measure can be a function of more

than one factor. Furthermore, we will assume that factor loadings  $\varphi$  will not depend on treatment  $D$ . This assumption implies that the effect of treatment on the measurements operates only through the latent factors and not the measurement system. This assumption can be easily tested estimating the measurement system separately for treatment and control groups and comparing the intercepts and slopes.

The factor structure is characterized by the means factors  $E[\theta_{d,t}^K], E[\theta_{d,t}^I], E[\theta_{d,t}^B], E[\theta_{d,t}^P]$  and the variance covariance matrix across factors, treatment arms and periods  $\Sigma_\theta$ .

The measurement system described above requires some conditions for identification

- The model is flexible to include several factors by measure. This is particularly useful for some developmental scales which are likely to be related to both cognitive and non-cognitive skills. However, we need to restrict that at least one measure per latent factor is exclusive to that factor.
- We set the factor scale, which means we need to standardize the factor loading of one measure per latent factor and per period equal to 1, or  $\varphi_{1,t}^K = \varphi_{1,t}^I = \varphi_{1,t}^B = \varphi_{1,t}^P = 1$ .
- We set the factor location by setting the intercept of the first measures to zero,  $v_{1,t}^K = v_{1,t}^I = v_{1,t}^B = v_{1,t}^P = 0$
- The measurement error  $\eta_{j,t}^K$  is a mean-zero error term independent of the latent factors and of each other,  $E[\eta_{j,t}^K] = E[\eta_{j,t}^I] = E[\eta_{j,t}^B] = E[\eta_{j,t}^P] = 0$

We proceed in several steps:

1. Identification of factor means: This is achieved by using the factor locations of the first measure and the independence of measurement error

$$\begin{aligned} E[M_{d,1,t}^K] &= E[\theta_{d,t}^K] \\ E[M_{d,1,t}^I] &= E[\theta_{d,t}^I] \\ E[M_{d,1,t}^B] &= E[\theta_{d,t}^B] \\ E[M_{d,1,t}^P] &= E[\theta_{d,t}^P] \end{aligned}$$

2. Identification of factor covariances: We use the covariances of the measurement system. For example, let's assume for now that each measurement is related to only one factor, and we want to identify the factor loadings for child skills



$\varphi_{j,t}^K$ . If we have three measures for child skill  $K$  then we can write

$$\begin{aligned} M_{d,1,t}^K &= v_{1,t}^K + \theta_{d,t}^K + \eta_{1,t}^K \\ M_{d,2,t}^K &= v_{2,t}^K + \varphi_{2,t}^K \theta_{d,t}^K + \eta_{2,t}^K \\ M_{d,3,t}^K &= v_{3,t}^K + \varphi_{3,t}^K \theta_{d,t}^K + \eta_{3,t}^K \end{aligned}$$

taking covariances we obtain

$$\text{cov}(M_{1,1,t}^K, M_{0,1,t}^K) = \text{cov}(\theta_{1,t}^K, \theta_{0,t}^K)$$

### 3. Identification of the factor loadings

$$\begin{aligned} \text{cov}(M_{1,1,t}^K, M_{0,2,t}^K) &= \varphi_{2,t}^K \text{cov}(\theta_{1,t}^K, \theta_{0,t}^K) \\ \varphi_{2,t}^K &= \frac{\text{cov}(M_{1,1,t}^K, M_{0,2,t}^K)}{\text{cov}(M_{1,1,t}^K, M_{0,1,t}^K)} \end{aligned}$$

in the same way we find

$$\varphi_{3,t}^K = \frac{\text{cov}(M_{1,1,t}^K, M_{0,3,t}^K)}{\text{cov}(M_{1,1,t}^K, M_{0,1,t}^K)}$$

### 4. Identification of the variance of each factor

$$\begin{aligned} \text{cov}(M_{1,1,t}^K, M_{1,2,t}^K) &= \varphi_{2,t}^K \text{var}(\theta_{1,t}^K) \\ \text{var}(\theta_{1,t}^K) &= \frac{\text{cov}(M_{1,1,t}^K, M_{1,2,t}^K)}{\text{cov}(M_{1,1,t}^K, M_{0,2,t}^K)} \text{cov}(M_{1,1,t}^K, M_{0,1,t}^K) \end{aligned}$$

in the same way

$$\begin{aligned} \text{cov}(M_{0,1,t}^K, M_{0,2,t}^K) &= \varphi_{2,t}^K \text{var}(\theta_{0,t}^K) \\ \text{var}(\theta_{0,t}^K) &= \frac{\text{cov}(M_{0,1,t}^K, M_{0,2,t}^K)}{\text{cov}(M_{1,1,t}^K, M_{0,2,t}^K)} \text{cov}(M_{1,1,t}^K, M_{0,1,t}^K) \end{aligned}$$

### 5. Identification of the variances of measurement error: Is achieved using the

factor loadings and variance of factors achieved in 3. and 4.

$$\text{var}(\eta_{j,t}^K) = \text{var}(M_{d,j,t}^K) - (\varphi_{j,t}^K)^2 \text{var}(\theta_{j,t}^K)$$

6. Identification of the measurement means: As we have identified the factor loadings and the mean factors, their identification is straightforward

$$v_{j,t}^K = E[M_{d,j,t}^K] - \varphi_{j,t}^K E[\theta_{j,t}^K]$$

7. Identification of the production function: We want to identify the parameters  $\tau_d$ ,  $\tau_d^I$ ,  $\tau_d^B$ ,  $\sigma_d$  and  $\alpha_d$  from the production function

$$\theta_{d,t+1}^K = \tau_d + \tau_d^I \theta_{d,t+1}^I + \tau_d^B \theta_{d,t+1}^B + \sigma_d \theta_t^K + \alpha_d \theta_t^P + \beta_d X_t + \varepsilon_{d,t+1}$$

Here we follow the same logic as in previous steps. Child skills at follow-up are also measured with error, therefore the identification of each parameter is achieved taking covariances between measurements of outputs and measurements of inputs. For example, to identify  $\tau_d^I$  we first compute

$$\text{cov}(\theta_{d,t+1}^K, M_{d,1,t+1}^I) = \tau_d^I \text{cov}(\theta_{d,t+1}^K, \theta_{d,t+1}^I)$$

where the covariance across latent factors has been achieved in 2. We can perform the same analysis for the remaining parameters of the production function.

## Estimation

The estimation procedure in a linear framework is a simple three-stage approach. In the first stage, we estimate the measurement system specified in equation 4.9. In our framework, the selection of factors is rather straightforward. We have defined two factors per child skills (cognitive and non-cognitive) both at baseline and at follow-up. Regarding parental investments, we will start simple by estimating only one factor for child stimulation. Provided our data variability is good enough, we will extend the analysis to two factors, one for cognitive stimulation and one for non-cognitive stimulation. Regarding parental beliefs, we will include both factors,

one for parenting styles, and one for perceived self-efficacy. Finally, our data of parental traits allows us to measure one factor for parental cognitive skills (based on measures of language and quantitative skills), and one factor for non-cognitive skills (based on personality traits and scales of stress and depression).

In the second stage, we use the set of measures, the factor loadings and the identified covariance matrix of measurement errors to compute unbiased estimates of the vector of latent factors for each individual. Heckman and Pinto (2013) show that the unbiased estimator of the vector of factors is given by

$$\theta_i = (\varphi' \Omega^{-1} \varphi)^{-1} \varphi^{-1} \Omega^{-1} M_i$$

where  $\theta_i$  is the vector of latent factor of dimension  $dim(\theta)$ ,  $M_i$  is the vector of stacked measures for each participant subtracting the intercepts, whose dimension is the sum across latent factors for all the available measures per factor;  $\varphi$  is the matrix of factor loadings, whose dimension is  $dim(M) \times dim(\theta)$ ; and  $\Omega_i$  is the covariance matrix of the measurement error.

Finally, in the third stage we use the estimated factor scores to estimate by using standard least squares the equation 4.8 governing the productivity of inputs, equation 4.6 determining inputs, and equation 4.7 estimating the linear technology 4.7.

#### 4.5.4 A Structural Model of Parental Investments in Children

The framework developed in section 4.5 provides a first look at the mechanisms underlying program impacts. Although it is a statistical framework, it is very much motivated by economic concerns, since what we essentially model a production function and a policy function for investments and beliefs. In this section we fully specify a model of parental investments in children where beliefs and expectations have a central role in generating the investment function as optimal behavior. If we can estimate such a model we can then use it to simulate new policies.

To be specific, we plan to estimate a model in the fashion of Cunha and Heckman (2007) and Cunha and Heckman (2006) in which parents make choices about consumption and child investments at different child ages, and as a result of those investments child cognitive and non-cognitive outcomes accumulate according to a technology of skill formation. However, we depart from the standard approach by explicitly incorporating our psychology-grounded measures of beliefs entering the

technology, and by using the elicited perceived returns as determinants of the heterogeneity in parental investments (Badev and Cunha (2012)). Building on Cunha, Elo, and Culhane (2013), the returns to child investments are uncertain to mothers. From the perspective of the mother, returns are a random variable, mothers have prior beliefs about them, and they are updated by a learning process after realizing the child development status in the next period. Mothers have prior perceptions about the mean and the variance of returns which are normally distributed, which lends itself to a standard updating process using Bayesian methods.

We consider two dimensions of parental investments, cognitive and non-cognitive stimulation. Non-cognitive aspects of investment are related to the behavioral management and socio-emotional stimulation, while cognitive aspects of investments are more related to learning activities and language stimulation. These two latent factors are extracted from the measurement system detailed in section 4.5.3.

## Preferences

At each age of the child  $a$  the household value household consumption,  $c_a$ , and child's cognitive and non-cognitive skills  $\theta_a^k$ , through a cobb-douglas utility function

$$u(c_a, \theta_a^C, \theta_a^N) = \alpha_1 \ln(c_a) + \alpha_2 \ln(\theta_a^C) + \alpha_3 \ln(\theta_a^N) \quad (4.10)$$

where the decision unit is the household as we are not able to identify time use of mothers and fathers in the data. Our data on investments is related to the principal caregiver, which in 95% of the cases is the mother.

In this specification, parental preferences for child human capital captured by parameters  $\alpha_2$  and  $\alpha_3$  can be a function of observed and unobserved heterogeneity. Including unobserved heterogeneity in preferences, and as discussed below, in expectations, makes identification very difficult. But we discuss our strategy to solve for this issue below. Denoting  $b_a$  the household's assets,  $\theta_{C,a}^I$  cognitive stimulation and  $\theta_{N,a}^I$  non-cognitive stimulation, the budget constraints is defined by

$$b_{a+1} = (1 + r)b_a + y_a - c_a + \pi_C \theta_{C,a}^I + \pi_N \theta_{N,a}^I \quad (4.11)$$

where household income is stochastic and follows an AR(1) process

$$y_a = \rho y_{a-1} + \varepsilon_a$$

$$\varepsilon_a \sim N(0, \sigma_\varepsilon^2)$$

## The technology of skill formation

Child's cognitive and non-cognitive skills accumulate according to a Cobb-Douglas technology of skill formation which is estimated separately for both treatment and control groups. In further, we omit the index for treatment to simplify notation.

$$\ln(\theta_{a+1}^k) = \sigma_1^k \ln(\theta_a^C) + \sigma_2^k \ln(\theta_a^N) + \tau_1^k \ln(\theta_{C,a}^I) + \tau_2^k \ln(\theta_{N,a}^I) + \eta_a^k \quad (4.12)$$

The main difference with equation 4.7, is that now we depart from a linear framework and we estimate a more flexible specification for the technology. We now need to determine how our psychology-grounded measures of parental beliefs,  $\theta^B$ , and parental traits  $\theta^P$ , enter the model. We allow them to enter the technology by affecting both the level of investments and the productivity of such investments. We adopt a linear policy function for investments varying with beliefs, parental traits and other observables characteristics like

$$\theta_{k,a}^I = \beta_{k,0} + \beta_{k,1} \theta_a^P + \beta_{k,2} \theta_a^B + \beta_{k,3} X_a + \nu_{k,a} \quad (4.13)$$

However, parental beliefs and other pre-treatment variables will also determine the productivity of the stock of skills and the productivity of investments through the following equations

$$\begin{aligned} \sigma_1^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\sigma_1} + \zeta_{k,1}^{\sigma_1} \theta_a^B + \zeta_{k,2}^{\sigma_1} \theta_a^P + \zeta_{k,3}^{\sigma_1} Z_a \\ \sigma_2^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\sigma_2} + \zeta_{k,1}^{\sigma_2} \theta_a^B + \zeta_{k,2}^{\sigma_2} \theta_a^P + \zeta_{k,3}^{\sigma_2} Z_a \\ \tau_1^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\tau_1} + \zeta_{k,1}^{\tau_1} \theta_a^B + \zeta_{k,2}^{\tau_1} \theta_a^P + \zeta_{k,3}^{\tau_1} Z_a \\ \tau_2^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\tau_2} + \zeta_{k,1}^{\tau_2} \theta_a^B + \zeta_{k,2}^{\tau_2} \theta_a^P + \zeta_{k,3}^{\tau_2} Z_a \end{aligned} \quad (4.14)$$

where again  $Z$  and  $X$  can have common variables.

In the framework laid out so far, the only sources of uncertainty are given by income shocks and shocks to technology inputs. This means we are assuming that the production function is fixed, i.e., parents know the returns to the relevant inputs. This is a strong assumption, as parents usually take investment decisions with certain level of uncertainty about the returns. It is therefore more realistic to assume that

the returns are probabilistic and parents learn about them as they observe the outcomes of their previous investment decisions.

We model this in the following way. Parents have prior perceptions about the parameters of the production function, which are probabilistic. In particular, we collect data on subjective probabilities they put on the returns to cognitive and non-cognitive stimulation  $\tau_1^k$  and  $\tau_2^k$ , so they have certain probabilistic distribution

$$f(\tau_1^k, \tau_2^k | \theta^B, \theta_a^P, Z)$$

where

$$\begin{aligned}\tau_1^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\tau_1} + \zeta_{k,1}^{\tau_1} \theta_a^B + \zeta_{k,2}^{\tau_1} \theta_a^P + \zeta_{k,3}^{\tau_1} Z_a + \epsilon_{k,a}^{\tau_1} \\ \tau_2^k(\theta_a^B, \theta_a^P, Z) &= \zeta_{k,0}^{\tau_2} + \zeta_{k,1}^{\tau_2} \theta_a^B + \zeta_{k,2}^{\tau_2} \theta_a^P + \zeta_{k,3}^{\tau_2} Z_a + \epsilon_{k,a}^{\tau_2}\end{aligned}$$

From either of these models we can derive reduced form relationships between investments and beliefs, which we can then match with the data to try to back out the parameters of the model. However, it is hard to separate unobserved heterogeneity in preferences from unobserved heterogeneity in perceptions, unless they can be assumed to be independent. Furthermore, parental beliefs cannot affect preferences, but only the levels of investments and the perceptions about the parameters of the production function. This is itself a strong assumption.

In order to allow unobserved heterogeneity drive both preferences and expectations, we follow the approach of Badev and Cunha (2012) collecting expectations data directly, as described in the next section.

### Eliciting perceived returns to parental investments

We attempt to separately identify heterogeneity in preferences and expectations by collecting expectations data about the returns to investments. In practice, this is only going to be possible for few parameters, so we focus on the returns to cognitive and non-cognitive stimulation.

The idea to collect data that allows us to directly recover:

$$f(\tau_1^k, \tau_2^k | \theta^B, \theta_a^P, Z)$$

Then, using choices and budget constraints, one could in principle back out the parameters of the utility function. But again one needs assumptions, if there are unobserved input about which one is not expressing expectations on, or unobserved

endowments. So this data will help a lot, but we still need some orthogonality assumptions.

The target of collecting data about the vector of perceived returns to investments  $\{\tau_1^k, \tau_2^k\}$  is to identify a prior distribution that then we can use as a starting point for the learning process. We assume the vector of priors is normally distributed

$$\begin{bmatrix} \tau_1^k \\ \tau_2^k \end{bmatrix} \sim N \left( \begin{matrix} \mu_{0,1}^k \\ \mu_{0,2}^k \end{matrix}, \Sigma_0^k \right) \quad (4.15)$$

The prior distribution is then updated using standard bayesian rules for the mean and variance of normally distributed variables.

We now proceed to describe how the instrument to elicit perceived returns helps us identify the prior distributions of the returns to cognitive and non-cognitive stimulation.

We start by noting that the mother doesn't know the technology, so she thinks this is

$$E[\ln(\theta_{a+1}^k)] = \sigma_1^k \ln(\theta_a^C) + \sigma_2^k \ln(\theta_a^N) + \mu_{0,1}^k \ln(\theta_{C,a}^I) + \mu_{0,2}^k \ln(\theta_{N,a}^I) \quad (4.16)$$

If we identify the prior distribution, we will also be able to estimate a production function in which the mother learns about the returns to investments. In a setup with one investment and one developmental measure, Cunha, Elo, and Culhane (2013) show they can identify the means and variances characterizing the distribution of returns by collecting data of expected developmental milestones in two investment scenarios and two levels of initial endowments (health at birth). This approach is useful if one is interested in estimating a production function with only one measure of investment.

In our model, we are interested in cognitive and non-cognitive investments so we construct four different home environments, combining high/low cognitive stimulation with high/low application of positive disciplinary strategies. To characterize a rich/poor home environment for cognitive stimulation, we analyzed the HOME test and analyzed parental behaviors belonging to the p75/p25 of the overall score. We followed the same strategy to characterize rich/poor home environments for positive disciplinary strategies, analyzing the scale Parenting Behavior Checklist, sub-scale Discipline.

For each of the four home environments, we then ask the mother to report the

minimum and the maximum age in which a child of certain age should be able to achieve successfully a certain lists of tasks, which were carefully taken from the list of tasks that a child of that age should be able to complete successfully 12 months after in the language and the socio-emotional tests. The idea is to be able to anchor maternal beliefs about child development to the actual development of the child we will observe in the follow-up survey.

For example, suppose we ask the mother: “If the child is currently between 12 and 17 months-old, what do you think is the youngest age and the oldest age a child learns to recognize body parts such as nose, eyes, feet, hands, mouth, etc?”. We ask the mother to answer the age range for each of the four home environments. Suppose the mother response to this task for one particular home environment (for example rich/cognitive and rich/positive discipline) is {18, 28}. If we assume a uniform distribution between the min/max age, then we should conclude that the probability that the child completes successfully that task at age 24 is 0.5.

For illustration of how the instrument identifies the distribution of returns, suppose  $[\bar{\theta}_C^I, \underline{\theta}_C^I]$  are the high/low cognitive stimulation scenarios and  $[\bar{\theta}_N^I, \underline{\theta}_N^I]$  the high/low non-cognitive stimulation scenarios. From mothers responses about age-specific developmental tasks we can recover the expected outcomes under different combinations of scenarios  $E[\theta_a^k | \bar{\theta}_C^I, \bar{\theta}_N^I]$ ,  $E[\theta_a^k | \bar{\theta}_C^I, \underline{\theta}_N^I]$ ,  $E[\theta_a^k | \underline{\theta}_C^I, \bar{\theta}_N^I]$ , and  $E[\theta_a^k | \underline{\theta}_C^I, \underline{\theta}_N^I]$ . Using the technology specification we can write for example

$$E[\theta_a^k | \bar{\theta}_C^I, \bar{\theta}_N^I] = \sigma_1^k \ln(\theta_a^C) + \sigma_2^k \ln(\theta_a^N) + \mu_{0,1}^k \ln(\bar{\theta}_{C,a}^I) + \mu_{0,2}^k \ln(\bar{\theta}_{N,a}^I) \quad (4.17)$$

This expectation can be recovered from our elicited subjective probabilities by using IRT methods. The idea of the methodology is to provide a map between an underlying factor and the probability with which some particular values of that trait occur. Here we take advantage of the fact that mothers have provided us subjective probabilities about developmental tasks that we actually assess in children of the same age and similar household characteristics. Therefore we can map those probabilities with expectations in equation 4.17 in the way it is shown in Figure 4.2

Once the expectations have been recovered, it can be shown that

$$\mu_{0,2}^k = \frac{E[\theta_a^k | \bar{\theta}_C^I, \bar{\theta}_N^I] - E[\theta_a^k | \bar{\theta}_C^I, \underline{\theta}_N^I]}{\ln(\bar{\theta}_N^I) - \ln(\underline{\theta}_N^I)}$$



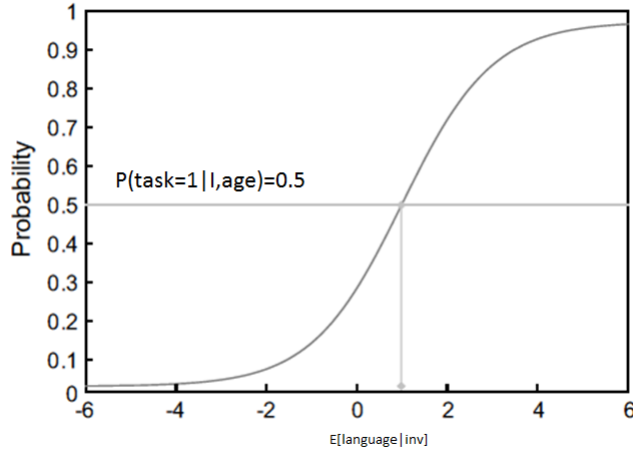


Figure 4.2: Mapping Subjective Probabilities to Conditional expectations through IRT models

and

$$\mu_{0,1}^k = \frac{E[\theta_a^k | \bar{\theta}_C^I, \bar{\theta}_N^I] - E[\theta_a^k | \underline{\theta}_C^I, \bar{\theta}_N^I]}{\ln(\bar{\theta}_C^I) - \ln(\underline{\theta}_C^I)}$$

## 4.6 Baseline Data

In this final section we discuss the main insights of the baseline data, describing the target population, the balance of the sample, the socio-economic gradients for the main inputs and outputs, and we provide multivariate correlations setting the ground for the empirical analysis when the follow-up data is available (October 2014).

### 4.6.1 Target Population

NEP serves the poorest socio-economic groups who are attended in the public health system. The average monthly income per-capita of the household in our baseline sample is Ch\$85.960 (US\$ 171.9). In contrast, the monthly average household income per capita of the Chilean population in 2011 was Ch\$ 243.670 (US\$ 487.3), according to the CASEN 2011 survey. NEP participants are much poorer than the Chilean population and than the standard user of the public health system, which makes this evaluation more interesting from a developmental perspective. In fact, 52.1% of our sample are in the first quantile of the national income distribution, and only 1% in the fifth quantile. Figure 4.3 shows the distribution of incomes.

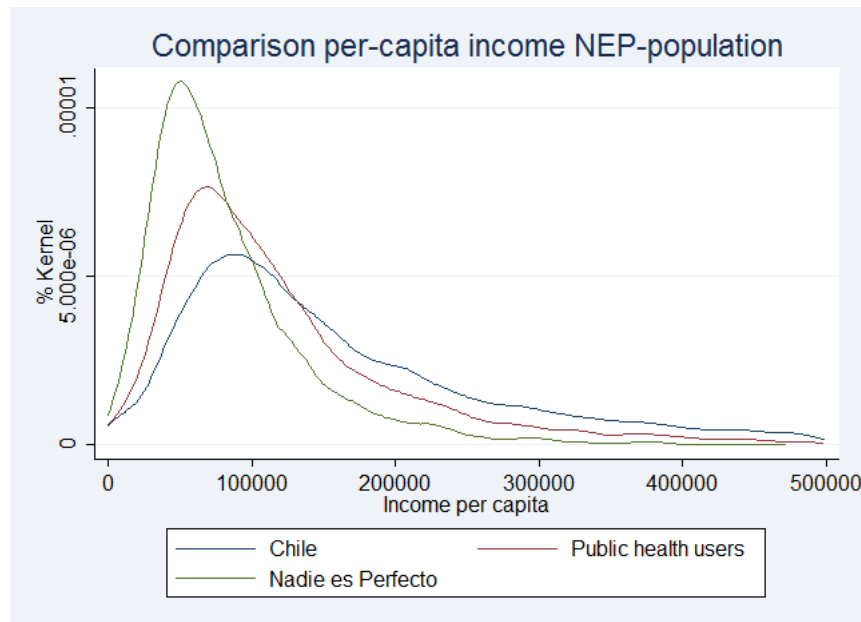


Figure 4.3: Income per capita in CLP: i) National, ii) National users of the public system, iii) NEP

## 4.6.2 Data descriptives

### Sociodemographics

Table 4.5 presents the main descriptive statistics and sample balance for the 2,916 principal caregivers and 3,597 children participating in the evaluation. Among them, 53.38% are males, and the average age was 27.96 months. The majority of children in the study were below 2 years old (47.4%), however the age distribution is widespread. We don't find significant differences across groups nor in gender or in the order of the child among all siblings, but for children below 12 months old in the Control group. Caregivers are mostly mothers (94%), followed by grandmothers. We find a very low presence of the father in our sample (1.5%), which is consistent with the data from the program. The average age of caregivers is 28.94 years old. The biggest proportion of our sample of caregivers are between 21 and 30 years old. Regarding education, roughly 60% of our sample finish secondary school, and 16% has some level of College. As discussed above, the representative family in our sample is very poor, with an average per-capita household income of 86 US\$/month. Participant families are in a large proportion bi-parental non-extended (presence of the father and the mother, with no other adults at home). 10% of the sample are mono-parental non-extended (single mothers with no presence of other adults).

It's important to notice that except from the youngest age group of children,

our sample is balanced across treatment groups for all the main socio-demographic characteristics of participants.

Sample Socio-demographics and Balance	Observations	Control	NEP Basico	NEP Intensivo
<b>Children</b>	<b>3597</b>			
Males		0.54	0.53	0.54
Age (months)				
0-12		0.28*	0.25	0.26
13-24		0.21	0.25	0.23
25-36		0.16	0.19	0.17
37-48		0.18	0.16	0.17
49-60		0.09	0.10	0.11
61-72		0.07	0.06	0.06
<b>Principal Caregiver</b>	<b>2916</b>			
Relationship with the child (proportions)				
Mother		0.95	0.95	0.95
Grandmother		0.04	0.03	0.03
Father		0.01	0.01	0.01
Age (years)				
13-20		0.13	0.15	0.14
21-30		0.49	0.49	0.49
31-40		0.30	0.27	0.28
41-50		0.07	0.08	0.08
>51		0.02	0.02	0.02
Education (proportions)				
Primary		0.20	0.23	0.20
High School Dropout		0.17	0.17	0.16
High School Degree		0.47	0.44	0.47
College		0.16	0.16	0.17
<b>Family Demographics</b>	<b>2916</b>			
HH Per-capita Income (M\$/month)		87.26	83.72	86.89
Family type (proportions)				
Monoparental - No extended		0.10	0.11	0.11
Monoparental - Extended		0.32	0.31	0.32
Biparental - No extended		0.41	0.39	0.38
Biparental - Extended		0.17	0.19	0.19

Table 4.5: Descriptive Statistics and Sample Balance Socio-demographic variables

Note: (\*) significant differences across treatment groups at 10%; (\*\*) significant differences at 5%

## Beliefs and Investments

Table 4.6 describes basic descriptive statistics and sample balance for parental beliefs and investments. The IRT estimates of the scale Ideas About Parenting measuring parenting styles do not show significant differences across treatment arms. We also do not find significant differences in raw scores across groups in parental perceived

self-efficacy, nor in perceived social support. Regarding investments, IRT estimates of the HOME scale do not show significant differences across groups. The same is true for the PBC Nurturing and Discipline scales.

<b>Parental Beliefs</b>	<b>Total</b>	<b>Control</b>	<b>NEP Basico</b>	<b>NEP Intensivo</b>	<b>p</b>
<b>Parenting Styles: IRT scores</b>					
<b>Autorithative</b>	0.39	0.41	0.37	0.39	0.550
<b>Autoritarian</b>	-0.28	-0.27	-0.29	-0.28	0.877
<b>Permissive</b>	-0.53	-0.54	-0.51	-0.54	0.324
<b>Perceived Self-efficacy: raw scores</b>	63.99	63.88	63.84	64.27	0.633
<b>Perceived Social Support: raw scores</b>	3.45	3.47	3.43	3.46	0.814
<b>Parental Practices</b>					
<b>Home test: IRT score</b>	0.79	0.81	0.77	0.79	0.393
<b>PBC Nurturing: raw scores</b>	4.00	4.00	3.99	4.02	0.471
<b>PBC Discipline: raw scores</b>	2.72	2.73	2.73	2.69	0.256

Table 4.6: Descriptive Statistics and Sample Balance Beliefs and Investments  
 Note: (\*) significant differences across treatment groups at 10%; (\*\*) significant differences at 5%

## Child outcomes

Finally, Table 4.7 describes child outcomes and our baseline sample balance. The first panel shows child's performance in language development measured with the scale PLSIV. A 70% of our sample between 3 months and 5 years old are diagnosed in the normal range, while 4,5% are diagnosed with clinical delays. We do not find significant differences across groups in any sub-scale.

The second panel describes executive functions measures, which are defined as the cognitive aspects of self-control like working memory, inhibitory control and attention shifting. The Dimensional Card Sort scale (DCCS) measures executive functions performance in children older than 24 months old. In the test, if the child doesn't pass the first stage, she cannot be evaluated, which means that her performance is too low to be measured by the scale. If the child passes the first stage, she is evaluated as "Normal" if she completes the task, or "Altered" if she leaves the task incomplete. The table shows the proportion of children with "Altered" results out of those who passed the first stage. We din't find any significant differences in the diagnostic across groups, except for the age groups 24-25 and 60-72 months.

The last two panels shows our measures of non-cognitive development. The positive aspects of socio-emotional development are measured with the Battelle scale. In our sample, 73,3% of the children 0 to 5 years old showed a normal performance. Some negative aspects of emotional development are captured by the CBCL scale measuring behavioral problems with pairs and adults, and it is applied to children between 18 months and 5 years old. In our sample, more than 28% of the sample shows some mild or severe level of alteration. We do not find any significant differences in scores or diagnostic across groups in these two dimensions.

Child Outcomes	Control	NEP Basico	NEP Intensivo
<b>Language Development (proportions)</b>			
Clinical delay	0.05	0.04	0.04
Risk	0.25	0.24	0.27
Normal	0.70	0.72	0.69
<b>Executive Functions (proportion with delays)</b>			
24-35 meses	0.47*	0.28*	0.42
36-47 meses	0.27	0.30	0.29
48-59 meses	0.15	0.25	0.17
60-72 meses	0.13*	0.07	0.16
<b>Socio-emotional Development</b>			
Clinical range	0.07	0.07	0.07
Borderline	0.19	0.19	0.20
Normal	0.74	0.73	0.73
<b>Disruptive Behavior</b>			
Clinical range	0.16	0.14	0.14
Borderline	0.13	0.16	0.14
Normal	0.72	0.71	0.72

Table 4.7: Descriptive Statistics and Sample balance child outcomes

Note: (\*) significant differences across treatment groups at 10%; (\*\*) significant differences at 5%

### 4.6.3 Socioeconomic Gradients

The importance of socio-economic gradients in explaining early gaps in cognitive and non-cognitive child outcomes in developing countries is well documented in the literature of child development (see for example Fernald, Weber, Galasso, and Ratsifandrihamanana (2011) and Schady, Behrman, Araujo, Azuero, Bernal, Bravo, Lopez-Boo, Macours, Marshall, Paxson, and Others (2014)). By showing that these socio-economic gradients also occur at the level of parental investments, some authors have emphasized the role of credit constraints driving sub-optimal parental investments in children. Yet there is much about the role of family background that we still do not understand, in particular the role of parental beliefs and attitudes.

In this section, we provide evidence that parental beliefs and expectations are potentially at the root of the sharp socioeconomic differences found in outcomes and investments.

Figure 4.4 documents SES gradients in child's cognitive skills. Panel a) examines language development. Children of more educated mothers obtain significantly higher standardized scores in both receptive and expressive language. If we plot the same figure but using parental income quintiles instead, the graph shows exactly the same pattern. Panel b) presents the SES gradients for child's performance in executive functions, where the y-axis measures the number of unsuccessful trials in the task before performing it correctly.

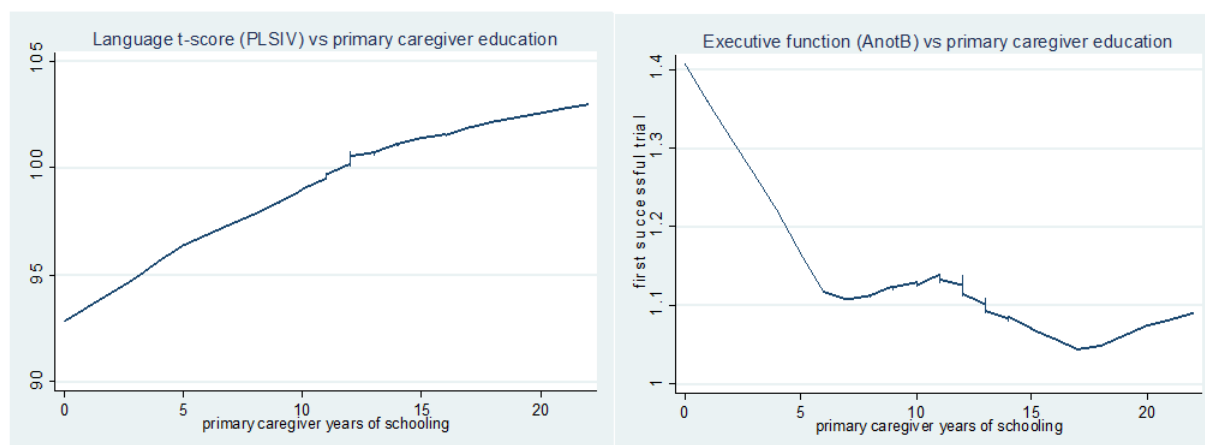


Figure 4.4: SES gradients in cognitive skills; a) Language; b) Executive functions

Figure 4.5 documents SES in non-cognitive skills. Panel a) presents the results for maladaptive behavior measured with the CBCL scale, which basically test the negative aspects of socio-emotional development. Children of more educated mothers present less behavioral problem, trend that is repeated among all the sub-scales of the test (somatization, sleeping problems, attention disorders, emotional reactive, etc). Panel b) presents the gradients for the Battelle test, which measures the positive aspects of socio-emotional development. Again, children of more educated mothers show larger social abilities in dimensions like interactions with adults, interactions with peers and social adaptation.

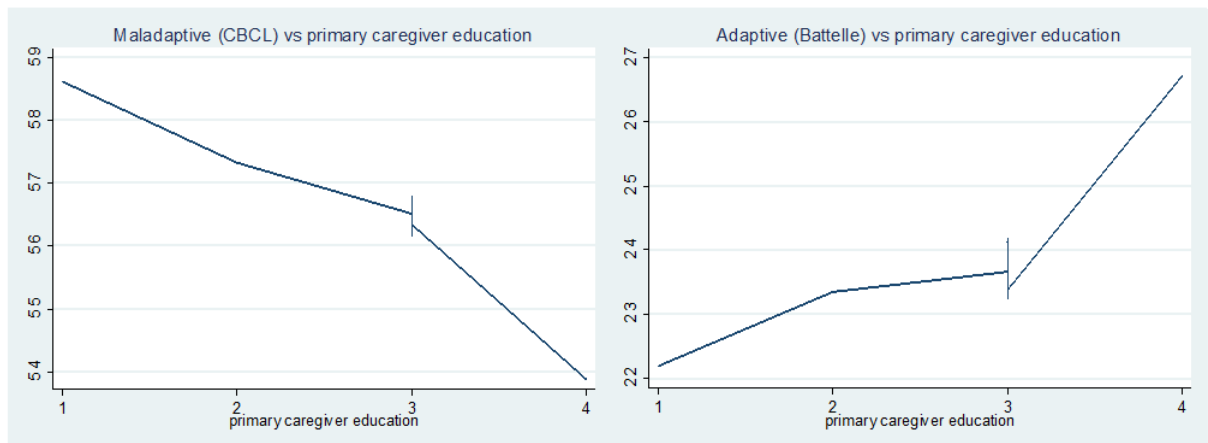


Figure 4.5: SES gradients in non-cognitive skills

Figure 4.6 confirms previous findings of sharp socioeconomic gradients in cognitive and non-cognitive parental investments (Moon (2010)). Panel a) presents describes the gradients for non-cognitive stimulation strategies. The Parenting Behavior Checklist sub-scale Nurturing, which measures socio-emotional stimulation and positive reinforcing of desirable child behavior, is strongly positively correlated to maternal education. Instead, the sub-scale of Discipline, which measures the use of harsh disciplinary strategies to correct for undesirable disruptive behavior of the child, is negatively associated to maternal education. Panel b) presents the SES gradients for cognitive stimulation strategies measured with the Family Care Indicators scale. The test, which includes measures of the number of books and magazines at home, the time spent with children in learning and play activities, and the variety of materials for learning and play, is positively correlated to maternal education.

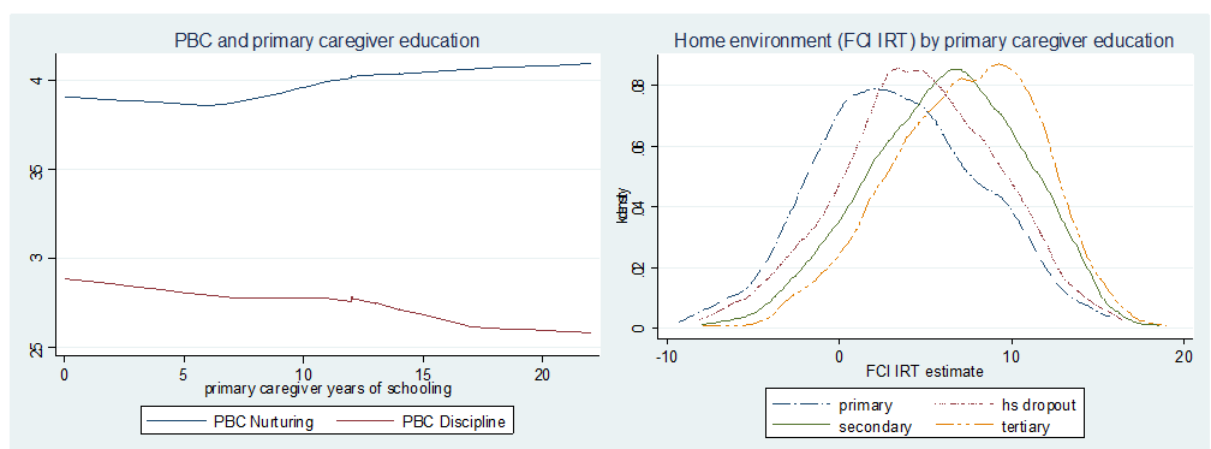


Figure 4.6: SES in parental investments a) Non-cognitive stimulation; b) Cognitive stimulation

Finally, Figure 4.7 present the SES gradients for our psychology-grounded measures of beliefs. Panel a) shows that more educated parents tend to see their role as parents less associated to authoritarian and permissive styles, and more associated to the authoritative style. Recall that the latter parenting style is associated to positive parenting and combines warmth and structure in a balanced way. The patterns are identical if the scales are plot against household income. Panel b) presents the SES gradients for perceived self-efficacy. More educated mothers also perceive themselves more competent and have higher self-esteem in the child-rearing task.

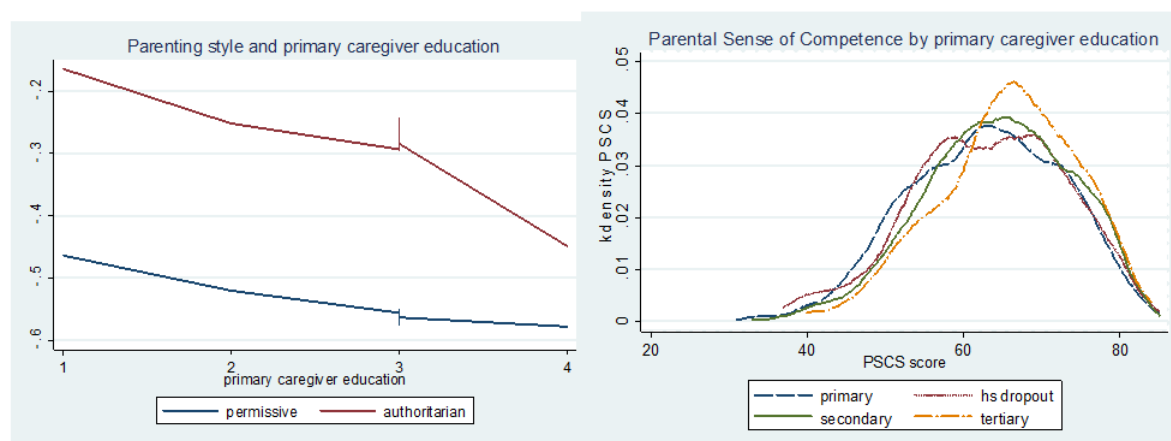


Figure 4.7: SES gradients in parental beliefs a) parenting styles; b) perceived self-efficacy

#### 4.6.4 Stylized Facts

The baseline data suggests that some strong associations between parental cognitive and non-cognitive stimulation and our psychology-grounded measures of parental beliefs. Note that we cannot include data on elicited perceived returns to investments because this data is only collected at follow-up. Table Regressions Parental Investments on Psychology-grounded parental beliefs presents multivariate regressions of error-free IRT estimates of investments on error-free estimates of parental beliefs. We control for family socioeconomic background and demographics, family composition and caregiver mental health. Some interesting results can be noticed. The authoritative style, which is associated to positive parenting, is positively associated to cognitive stimulation (home environment scales) and socio-emotional stimulation (PBC Nurturing scale). The authoritarian style is negatively associated to both cognitive and socio-emotional stimulation, but is positively associated to the use of harsh disciplinary strategies. Regarding perceived self-efficacy, correlations with



investments go in the proposed directions: parents who perceive themselves more competent or have higher self-esteem in the parental task also stimulate more their children and apply less harsh disciplinary strategies.

Dependent Variable: Parental Cognitive and Non-cognitive stimulation			
Dependent	Home Environment (FCI - IRT estimates)	Nurturing (PBC)	Discipline (PBC)
<b>Beliefs</b>			
Parenting Style			
Authoritarian	-0.475*** (0.12)	-0.041*** (0.01)	0.082*** (0.02)
Authoritative	0.417*** (0.11)	0.076*** (0.01)	0.079*** (0.01)
Permissive	-0.228 (0.15)	0.067*** (0.01)	0.020 (0.02)
Self-competence	0.030*** (0.01)	0.007*** (0.00)	-0.013*** (0.00)
Perceived social support	0.220*** (0.05)	0.032*** (0.01)	0.012* (0.01)
<b>Caregiver Mental health</b>			
Stress	-0.019** (0.01)	-0.001 (0.00)	0.005*** (0.00)
Depression	-0.002 (0.01)	0.002*** (0.00)	0.002 (0.00)
Constant	-4.609*** (1.07)	2.977*** (0.11)	2.965*** (0.13)
R2	0.23	0.22	0.23
N	3,252	3,257	3,256
Controls	Caregiver Education and Age; Household Income; Father Education; Family Structure (Single mother, extended family, older siblings)		

Table 4.8: Regressions Parental Investments on Psychology-grounded parental beliefs

Table Regressions Child Outcomes on parental investments and parental beliefs presents correlations between different dimensions of child outcomes and parental investments (specification 1), and the same associations but including our psychology-grounded measures of parental beliefs (specification 2). Most of the associations go in the directions proposed by our model of change. For example, cognitive stimulation (FCI scale) is positively correlated to language, executive functions and socio-emotional development, while it is not associated to maladaptive behavior (positive scores of this scales are related to worse outcome). Socio-emotional stimulation (PBC Nurturing scale) is positively correlated to socio-emotional development and negatively correlated to maladaptive behavior, correlations that remain stronger even after including parental beliefs. Disciplinary strategies are positively correlated to maladaptive behavior, even though they are also positively correlated to language

development and socio-emotional development. Analysis from a factor model shows that the reason for this result is because this scale actually entails two latent factors, one related to positive disciplinary strategies and another related to punitive disciplinary strategies. The first factor correlates to language and socio-emotional development, while the second factor correlates to maladaptive behavior.

Tables Regressions Parental Investments on Psychology-grounded parental beliefs and Regressions Child Outcomes on parental investments and parental beliefs shed light about the role of parental beliefs in child development. Beliefs are strongly correlated to parental investments, supporting the idea of the indirect effect of beliefs in child outcomes through investments discussed in section Decomposing treatment effects. But after controlling for parental investments, beliefs also have significant effects in child outcomes, suggesting that there might be a direct effect for example through changes in the productivity of investments. Moreover, the correlations below go all in the proposed directions: authoritative parents have children performing significantly better in language, executive functions and socio-emotional development. Instead, children of authoritarian show lower language development, and children of permissive parents have higher levels of behavioral problems. Finally, perceived self-competence is strongly associated to socio-emotional development and perceived social support is positively associated to executive functions and socio-emotional skills.

Dependent Variable: Child Development Outcomes								
	Language (PLSIV)		EF >24months (DCCS)		Maladaptive behavior (Achenbach)		Socioemotional (Battelle)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
<b>Parental Investments</b>								
Home environment (FCI - IRT estimates)	0.484*** (0.11)	0.461*** (0.11)	0.025*** (0.01)	0.024*** (0.01)	-0.153 (0.10)	-0.138 (0.10)	0.487*** (0.07)	0.486*** (0.07)
Nurturing (PBC)	1.302* (0.74)	1.064 (0.76)	0.014 (0.05)	0.002 (0.05)	-1.822*** (0.62)	-2.002*** (0.63)	1.698*** (0.48)	0.894* (0.49)
Discipline (PBC)	1.186** (0.59)	1.249** (0.60)	-0.022 (0.04)	-0.013 (0.04)	6.068*** (0.52)	5.634*** (0.54)	0.520 (0.37)	0.725* (0.38)
<b>Beliefs</b>								
Parenting Style								
Authoritarian		-1.540*** (0.50)		-0.020 (0.03)		0.565 (0.42)		-0.231 (0.32)
Authoritative		1.777*** (0.46)		0.082*** (0.03)		0.031 (0.39)		0.948*** (0.30)
Permissive		-0.778 (0.59)		-0.102*** (0.04)		2.010*** (0.49)		0.574 (0.39)
Self-competence		-0.016 (0.04)		0.003 (0.00)		-0.054 (0.04)		0.133*** (0.03)
Perceived Social Support		0.420 (0.22)		0.035*** (0.01)		0.062 (0.06)		0.238** (0.13)
Constant	73.939*** (4.27)	74.666*** (5.73)	1.665** (0.82)	1.446* (0.84)	37.868*** (3.381)	45.407*** (4.395)	3.279 (2.493)	-4.721 (3.318)
R2	0.15	0.17	0.30	0.31	0.18	0.20	0.18	0.19
N	2,901	2,901	1,695	1,695	2,082	2,082	3,192	3,192
Controls: Caregiver age and education, Household Incomes, Father Education, Family Structure, Child age and gender								

Table 4.9: Regressions Child Outcomes on parental investments and parental beliefs

## 4.7 Concluding remarks

Investments in human capital are central for economic growth and for the reduction of inequality. They are particularly relevant for improving the situation of the poor and breaking the intergenerational transmission of poverty. That said, the question of how best to foster human capital formation among the most disadvantaged is not an easy question to answer. There exists increasing evidence that high quality early childhood investments produce gains in child development that translate into improved long-run outcomes. However, much less is known about the channels through which these programs operate, and how high quality investments can be fostered in the family. Moreover, the provision of additional financial resources to poor families does not automatically translate into better development (e.g., Mayer (1997)). This demands a shift in research focus, from understanding the impacts of resources, to understanding the determinants of behaviors.

We attempt to address these questions by conducting the impact evaluation of a nationally-scaled center-based parenting program in Chile using Randomized Control Trials methods. The intervention is likely to change three dimensions of parental beliefs: beliefs about what is the best way to raise children, beliefs about how confident parents feel about the parenting task, and parental perception about the returns to investments in children. By collecting pre and post-treatment data from a sample of more than 3,000 households, we investigate the effects of these beliefs on parental investments and on child development.

Our project has several innovations. First, this is the first work collecting several measures of beliefs that complement each other, forming a much more complete set of beliefs than ever considered before. Second, we use the exogenous variation of beliefs provided by the intervention to investigate treatment effects controlling for measurement error. Third, we investigate the indirect effects of beliefs in outcomes by changing parental investments and the direct effects by changing the productivity of those investments under relatively weak assumptions. Fourth, we take advantage of the experimental data to separately identify heterogeneous preferences and expectations in the estimation of a model of parental investments in children in which we emphasize that parents are uncertain about the returns to investments in the technology of skill formation. Such a model will be potentially useful for the simulation of more cost-effective early child development policies.

Evidence from the baseline data is encouraging about the potential findings. First, the randomization worked across all our measurements. Second, evidence from the socio-economic gradients show that parental beliefs might be at the root of the well documented gradients in child development and parental investments, which we also find in our data. Third, the policy works with the most disadvantaged sectors of the Chilean population, so the potential returns of the intervention are big, and externally valid to the poorer segments of other countries. And finally, the multivariate associations are supportive of our modeling framework. Parental beliefs are strongly correlated (in the right directions) with parental investments, and parental investments strongly correlated to several dimensions of child outcomes, which is line with the suggested indirect effects of beliefs in outcomes. Moreover, we also find some evidence of a direct effect of beliefs in child outcomes, which is line with the idea that beliefs and attitudes also modify the productivity of such investments.

# Chapter 5

## A Collective Model of Labor Informality and Self-employment

### 5.1 Introduction

In this work in progress I sketch a model that studies comprehensively all the interacting incentives to work informally and in self-employment activities over the life-cycle in developing countries. I focus the analysis on four driving forces. The first one is comparative advantage, where I assess how sector-specific returns to human capital and preferences for job amenities jointly determine self-selection into schooling and into a particular sector of employment. The second one is the design of the welfare system. I assess how contributed-based pension and health systems and different conditional and unconditional welfare influence sorting into informal jobs or self-employment. Third, I study the extent to which informality and self-employment are the result of optimal collective decisions maximizing the output of home production, which could happen either because of these labor markets are more flexible and then more compatible with home production tasks like child-rearing, or simply because of by choosing different working sectors there is optimal intra-household risk sharing. Finally, I assess the importance of capital accumulation through asset holding for self-employment.

The modeling framework is a life-cycle model in which single-earner and double-earner households jointly decide schooling, labor supply and consumption. Individuals are heterogeneous in skills and preferences, feature that is modeled as permanent unobserved heterogeneity. Individuals start making decisions early in the life-cycle, so schooling is endogenous to sector-specific labor market expectations and potentially influenced by marriage prospects. The schooling decision involves to complete

secondary education or not, and to attend different types of post-secondary education, which are associated to substantial differences in labor market returns. Work participation involves four exclusive sectors: formal salaried employment, informal salaried employment, self-employment, and home production/non-participation. A key aspect of our model is that family composition is determined endogenously. Individuals get marriage according to an unobserved match quality preferences, which is adjusted dynamically throughout time as single individuals learn about their outside option.

This work in progress attempts to answer the general research question of why people sort into the informal sector and into self-employment. In particular, it will tackle the following specific questions:

1. Does comparative advantage, and in particular education, provide access to better jobs particularly in the formal sector?
2. Do individuals self-select into informality and self-employment based on unobserved skills and preferences?
3. How do the incentives provided by a fully-funded pension and health system affect informality and self-employment?
4. How do the incentives provided by welfare schemes like unemployment insurance and wage subsidies affect informality and self-employment?
5. What's the role of intra-household specialization in home and market activities and the role of flexibility in determining gender-specific work participation in the informal sector and in self-employment?
6. What's the role of asset accumulation in the size of self-employment?

As in Chapter 1, the model will be estimated using the same longitudinal household survey data from Chile, but now including both males and females. The survey retrieves detailed information on educational choices at the secondary and post-secondary levels, sector-specific labor supply, wages, family assets and debts, and household composition. Data on social security contributions and welfare benefits is merged from administrative records.

This work intends to contribute to the literature of informality and self-employment in four aspects. First, I disentangle informality from self-employment. Informal workers are salaried employees hired without a contract and who are not paid the

social security contributions (illegal informality), while the self-employed are independent workers or firm-owners who have the choice to contribute or not to the welfare system. These two types of labor have been traditionally pooled in the literature of labor informality but they differ in observed and unobserved characteristics and life-cycle participation show opposite patterns. Informal employees are usually young workers that accept jobs offers without a contract as an entry cost to labor markets, and as they gain experience, they move to formal jobs. Thus informality decreases over the life-cycle. Instead, the self-employed usually start working in the formal sector and as they gain experience and accumulate capital, start running their own businesses. Therefore, self-employment increases over the life-cycle. Self-employment is often associated to a choice based on entrepreneurship skills and capital accumulation (Bosch, Goni, and Maloney (2007); Bosch and Maloney (2007)) while salaried informality is often related to labor market segmentation.

Second, there is still little evidence on how schooling is associated to informality and self-employment in a dynamic context. Few papers have documented that education is a passport to better jobs in the formal sector using longitudinal data, but they either consider education as exogenous or they fail to control for unobserved factors interacting with the schooling decision. For example, returns to human capital in self-employment activities require taking into account the role of entrepreneurship abilities. Or there may be little scope for educational policies reducing informality if sector-specific labor supply is a choice based on unobserved skills. In this work, individuals can self-select into higher education and into different types of college based on sector-specific returns, there is an explicit role for unobserved skills linking schooling and sector-specific comparative advantage, and I allow for expectations about marriage to affect the schooling decision.

Third, I study labor informality and self-employment in the context of endogenous labor supply and family formation decisions, features that have not been studied together before but they are supported by data descriptives. While female labor supply is only 44% (compared to 73% for males), females tend to work more informally and self-employed than males at all education levels, suggesting an important degree of intra-household specialization of labor supply not only in the extensive margin, but also in the choice of a particular sector. Informal jobs and self-employment are often associated with more flexible working hours, which may make these jobs more compatible with home production. Therefore, the dynamics of family composition would be endogenously determined along labor supply and consumption/savings decisions. With regard to marriage, it will be assumed that in an efficient frame-

work single individuals match optimally with another single household according educational group and match quality preferences. Divorce is determined optimally comparing current situation with the outside option (Voena and Bayok (2014), Mazzocco, Ruiz, and Yamaguchi (2014), Ligon, Thomas, and Worrall (2002), Pistaferri, Meghir, Voena, and Low (2013)).

The final intended contribution is regarding the policy evaluation dimension. Such a model will allow me to perform the ex-ante and ex-post evaluation of several policies. One of them is the reform to pension system. Joubert (2010), Joubert and Todd (2011) and Attanasio, Meghir, and Otero (2011a) study a first reform implemented in 2008 but in their papers they do not account for the difference between self-employment and salaried informality, they disregard human capital accumulation, and the dynamics of family formation is purely exogenous. In this work I put all these elements together and therefore a comprehensive evaluation of the effects on informality of changes to the social security system can be made. Moreover, as the contributed-based pension system is obligatory for salaried workers but voluntary for the self-employed, I will be able to simulate a forthcoming reform starting in 2015 making the contributions compulsory for the self-employed, which potentially can change the incentives from the welfare system to informality and self-employment. A second important forthcoming reform is the reform to the post-secondary education system. College education is very expensive in Chile (45% of average annual wage rates) and there is a high degree of heterogeneity in returns to different types of degrees and types of institutions comparable to the US. A new reform starting in 2015 will make College education for free for the first three income quintiles of the income distribution affecting all types of College institutions. Such a sharp exogenous change in the direct costs to schooling is expected to produce a sharp change in the distribution of college enrollment and in labor market participation.

Regarding ex-post policy evaluation, I will be able to estimate the long-term effects of recently implemented reforms. First, the introduction of an unemployment insurance scheme in 2002, which is partially financed by workers, the employer and the Government. Workers can claim the subsidy after having worked formally and contributed to the system for 12 months (over the last 24 months). The long-term effects of the policy and the change of incentives for human capital accumulation have never been evaluated before. Second, the implementation of an employment subsidy for young workers introduced in 2008. This means-tested subsidy is offered to workers between 18 and 25 years old belonging to the first two income quintiles. The scheme operates like a standard tax credit scheme changing the slope of the



budget constraint at different income ranges. Bravo and Rau (2012) analyze the short-term effects of the subsidy using a RDD approach, but there are no studies analyzing the effects of those subsidies over the life-cycle. Finally, the introduction of a state guaranteed credit program (CAE) to finance college education. Rau, Rojas, and Urzúa (2013) use a structural approach the short-term effects on college participation and drop-out rates, but the long-term effects (particularly labor market outcomes) of private loans to college education have not yet been studied.

## 5.2 The Model

This is partial equilibrium life-cycle model with three stages: young life, adult life, and retirement. In the young life, individuals decide whether to invest in human capital or not by first attending High School or not, and after so choosing different types of College. Educational choices depend upon the cost of effort of education, monetary costs of schooling like tuition fees, and expectations about the future regarding private returns in the labor market, and marriage and fertility prospects. The cost of effort is driven by ability endowments which are modelled as a flexible distribution of unobserved discrete and finite types. The role of these unobserved types will be fundamental for our modeling framework: they will provide a direct link between abilities and skills driving sector-specific wage returns when people enter into the adult life.

If individuals drop schooling at any stage, they enter the adult life. At this stage, people optimally choose consumption and labor supply optimally, but they also may decide to form a new family by getting married. The labor supply decision involves allocating time to the home production of a public good, or to market production in one of three sectors: the formal sector, the informal sector, and self-employment activities. Individuals can choose to work in market production part-time or full-time, so they also spend time in home production. The choice of sector determines important sources of individual and household insurance. Formal jobs, unlike informal jobs and self-employment, requires contributing to the social security system and therefore individuals have access to the private health care system which is of better quality. Formal jobs also provide insurance against income shocks after retirement, as earners compulsory accumulate 10% of their gross salary into a pension savings account which they can use only upon retirement (65 for males, 60 for females). The self-employed can choose to contribute voluntarily either to their pension account or to the health system. Individuals can also smooth out income or unemployment

shocks through borrowing and saving. This channel may be particularly important for informal workers or the self-employed who face a larger volatility of income and employment. Private savings may also have the role of insuring households that have worked little in formal jobs and therefore will face low incomes after retirement. In the case of married couples, individuals will also have more flexibility to adapt their labor supply to the requirements of home production. For example, women can take part-time jobs or they can work informally or self-employed, jobs which are often associated to more flexibility and autonomy.

The dynamics of marriage and divorce is characterized by an intertemporal collective model with no commitment. Single individuals draw potential partners every period and can get married, but they cannot commit to the initial allocation of resources in the future. Instead, every period they can renegotiate the initial plan or get divorced if it is optimal for one or the two spouses. In this modelling framework, individual preferences are determined by individual consumption and leisure, by the public good collectively produced at home, and by the matching quality if they are married (Mazzocco, Ruiz, and Yamaguchi (2014)). After marriage, each subsequent period partners evaluate marriage continuation by comparing the individual value of marriage against the outside option of divorce. Matching quality is unobserved but it will be set at the beginning of the marriage depending on spouses unobserved types, and then it will evolve according to matching shocks. Married households will jointly decide consumption and labor supply efficiently in a cooperative framework, where each of the spouses has certain bargaining power. I follow the approach of Ligon, Thomas, and Worrall (2002), Voena and Bayok (2014) and Pistaferri, Meghir, Voena, and Low (2013) to determine the Pareto weights. At the moment of the match, they are found so that gains from marriage relative to the outside option for each spouse are equated. After marriage, weights will evolve so that gains from marriage are equally splitted, and if at some point there is no weight that make both spouses to prefer marriage than divorce, the marriage terminates.

At the retirement stage, which starts at age 60 for women and 65 for men and lasts until age 75 for both, single or married individuals are not allowed to work and they will consume the household incomes, which will be provided by the individual accumulated pension savings, by household assets, and by any sort of public transfers. Married couples are not allowed to get divorce at this stage.

### 5.2.1 Timing and Decisions

Individuals of both gender ( $j = m, f$ ) start making educational choices at age 14 based on expected returns, directly observed costs to schooling, and unobserved heterogeneity. Unobserved skills are modeled as initial skill endowments by including discrete and finite  $k$  unobserved types (Heckman and Singer, 1984). Years of secondary schooling will be pooled into the first period of the model, while from age 18 onwards each period of the model will represent one year in the life-cycle. While in secondary schooling, individuals do not need to fund their own education and consumption, but they have to fund post-secondary education with own income or student loans. People may achieve five schooling paths, denoted by  $s$ . They may be high school dropouts ( $s = 1$ ), obtain a high school diploma ( $s = 2$ ), and at age 18 to start three exclusive post-secondary careers: a technical degree in a Professional Institute ( $s = 3$ ), often lasting 3 years, a college degree in a high quality university (usually public) lasting 5 years ( $s = 4$ ), or a college degree in a low-quality (usually private) university ( $s = 4$ ).  $d_{s,t}^j$  equals 1 if individual achieves maximum schooling type  $s$ , which I denote by  $Ed^j$ .

Denoting the age of individuals by  $t^j$ , individuals meet a potential partner of the same unobserved type with certain probability and can get married, which happens if the single-earner value function is lower than the value of marriage. Individuals hold their own assets before marriage, while household assets after marriage are pooled individual assets. Two-earners household utility will be the weighted sum of individual utilities, where the Pareto weights are endogenously determined according to the difference between the value of staying in marriage and the outside option for each of the household members. In the same way, two-earners households can split up into single-earner households by comparing the benefits of staying married, which depend on the matching quality, with the value of becoming a single-earner. When this happens, assets are divided according to their bargaining power. Children arrive with exogenous probabilities which depends on spouses education, marital status, female age, and the presence of other children.

After dropping out school, single-earners and two-earner households start making consumption and labor supply decisions. Two-earners household decisions are taken according to spouses' bargaining power determined at the moment of matching. Labor supply choices are the following: individuals can work as a salaried employee in the formal sector ( $a = 1$ ), work as a salaried employee in the informal sector ( $a = 2$ ), be self-employed ( $a = 3$ ), or home production ( $a = 4$ ). Individual consumption is  $c_t^j$  and career choices  $d_{a,t}^j$  equals 1 if individual chooses alternative  $a$ . Moreover, part-

time jobs only in the formal salaried or in the informal salaried sectors is allowed for females, alternative denoted by the variable  $p_t^j$ . Every period an earner decide to work in sectors  $a = \{1, 2, 3\}$ , they accumulate one additional year of experience  $k_t^j$ , or half a year if they work part-time. If one earner decides to work in the formal salaried sector, a fixed amount of her wage ( $\phi = 10\%$ ) is saved in her pension account, money that can only be used for retirement. Ultimately, this is a human capital investment model, so the wage offer in a particular sector of the economy is the realization of a sector-specific technology of skill production function, in which each skill component is valued according to a sector-specific equilibrium rental price. In our model, the skill vector comprises education, sector-specific experience and innate ability.

### 5.2.2 The Young Life

At this stage individuals make educational choices based upon expected returns to schooling, marriage prospects and realized costs. Decisions are strictly individual, as I don't allow people to get married until they finish schooling or dropout. Monetary costs of education at secondary level are very low and we assume that the level of assets at age 14 is zero, so schooling participation at this stage depends on ability and family characteristics. Instead, post-secondary education starting at age  $t = 18$  is costly, so own consumption and fees should be funded with own assets (parental wealth) or student loans, and I restrict preference for leisure to behave as if college students worked full-time in the formal sector. Denote by  $\mu_k$  is initial skill endowment of individual type  $k$  at age 14, and  $z_i$  a set of observed socioeconomic background variables of the student's family (parents education and SES index). The flow utility of attending schooling level  $s$  which lasts  $\tau_s$  periods is

$$U_{s,k}^j = \begin{cases} \delta_{s,0,k}^j \mu_k + \delta_{s,1}^j z^j - \eta_s^j & \text{if } s = 2 \\ \frac{(c_s^j)^{1-\sigma}}{1-\sigma} \exp\{(\delta_{s,1}^j (d_{1,t}^j = 1))\} + \delta_{s,2}^j z^j + \delta_{s,0,k}^j \mu_k + \eta_s^j & \text{if } s = \{3, 4, 5\} \end{cases}$$

Where  $\delta_{s,0,k}^j$  is the cost of effort of education at level  $s$  for individual type  $k$ , and  $\eta_s^j$  is a random effort cost.

Let  $\Omega_1$  the state space at the beginning of secondary schooling years. Individuals choose either to finish high school or to drop out and start working in a particular sector of employment, or leisure/home production picking the choice that maximizes life-time utility, where the choice-specific value function for attending secondary schooling is

$$V_{s=2}^j(\Omega_{t=1}^j) = U_{s=2,k}^j + Emax \left[ V_{s=\{3,4,5\}}^j(\Omega_{t=2}^j \mid Ed^j = 2), V_{t=2}^{j,work}(\Omega_{t=2}^j \mid Ed^j = 1) \right]$$

where  $V_{s=\{3,4,5\}}^j(\Omega_{t=2}^j \mid Ed^j = 2)$  represents the value of attending any of the college choices given that the individual decided to finish secondary schooling, and  $V_{t=2}^{j,work}(\Omega_{t=2}^j \mid Ed^j = 1)$  is the value of any working alternative given that the individual decided to remain as a high school drop-out.

At age 18 the problem is slightly different, because individuals must fund education with own assets. Let  $\Omega_{t=2}^j$  the state space at the beginning of the period which includes initial assets and previous education level. I allow for College students to drop out education before they finish their degrees, and I also allow them to switch education to a different College type. The choice-specific value function of attending College types  $s = \{3, 4, 5\}$  is

$$V_{t=2,s}^j(\Omega_{t=2}^j) = \underset{c_{\tau s}}{Max} \{ U_{s,k}^j + \beta EV_{t+1}(\Omega_{t+1}^j \mid \Omega_t^j) \}$$

*s.to.*

$$A_{t+1}^j = R^{\tau s} A_t^j - c_t^j - F_s$$

where  $F_s$  are observed tuition fees for each type of post-secondary education, and  $A_t^j$  is the individual level of assets at a particular age.

### 5.2.3 The Adult Life

#### Preferences

Single individuals have preferences over own consumption and leisure, and they value the output of home production. Married individuals additionally value the quality of the matching. Utility is assumed to be separable in consumption, leisure or sector-specific dis-utility of work, the output of home production  $Q_t$ , and the matching quality  $(\theta_{t,k})$ , which is assumed to be type-specific. Home production is assumed to be a cobb-douglas production function of the number of children, time at home, and work participation in informal jobs or self-employed. In our data I cannot distinguish between leisure and actual home production. Under this specification, I

can capture the importance of informality and self-employment either in providing more flexibility and autonomy to maximize home production, and the dis-utility they create by restricting access to fringe benefits and the access to a better quality of health both attached to formal contracts. Moreover, individuals switch jobs across sectors face transition costs, reflecting search costs, the availability of networks, or psychological costs associated to start a job in new environments.

I model individual preferences as

$$\begin{aligned}
U_{t,k}^j(c_t^j, d_t^j, p_t^j, Q_{t,k}, \theta_{t,k}) &= \frac{(c_t^j)^{1-\sigma}}{1-\sigma} \exp\left\{\phi_1^j d_{1,t}^j + \phi_2^j d_{2,t}^j + \phi_3^j d_{3,t}^j\right\} + \alpha^j \log Q_t^j(n_t, d_t^j, p_t^j) + \theta_{t,k} 1[m_t = 1] \\
&+ \psi_1^j d_{1,t}^j + \psi_2^j d_{2,t}^j + \psi_3^j d_{3,t}^j + \varphi_1^j \left(d_{1,t}^j, d_{4,t-1}^j + d_{2,t}^j, d_{4,t-1}^j + d_{3,t}^j, d_{4,t-1}^j\right) \\
&+ \varphi_2^j \left(d_{1,t}^j, d_{2,t-1}^j + d_{1,t}^j, d_{3,t-1}^j\right) + \varphi_3^j \left(d_{2,t}^j, d_{1,t-1}^j + d_{3,t}^j, d_{1,t-1}^j\right)
\end{aligned}$$

The labor supply vector  $d_t^j$  summarizes the vector of dummies for the choice of sector  $d_t^j = \{d_{1,t}^j, d_{2,t}^j, d_{3,t}^j, d_{4,t}^j\}$  while the vector of part-time choices for females is summarized by  $p_t^f = \{p_{1,t}^f, p_{2,t}^f\}$ . Dis-utility of work in the informal sector and in self-employment is captured by  $\psi_2^j$  and  $\psi_3^j$ .  $\theta_{t,k}$  is the unobserved quality of the matching if they are married. I restrict spouses of the same unobserved types to form a family, so this parameter is type-specific and evolves dynamically as

$$\theta_{t,k} = \theta_{t-1,k} + \xi_{t,k} \sim N(0, \Sigma_k)$$

Home production will follow a Cobb-Douglas production function depending on the number of children  $n_t$ , and I assume that households benefit if spouses spend more time at home. Additionally, I assume that self-employment can have an additional effect in home production as this type of labor gives more time flexibility to invest in home activities that can be combined with market production. The log linear household good production function is

$$Q_t^j = (1 + n_t)^{\gamma_1} (1 + d_{4,t}^j)^{\gamma_2 + \gamma_3 n} (1 + d_{3,t}^j)^{\gamma_4 + \gamma_5 n} (1 + p_t^f)^{\gamma_5 + \gamma_6 n}$$

For married couples the production function, I assume that male and female times invested in home production are perfectly substitutable.

$$Q_t = (1 + n_t)^{\gamma_1} (1 + d_{4,t}^m + d_{4,t}^f)^{\gamma_2 + \gamma_3 n} (1 + d_{3,t}^m + d_{3,t}^f)^{\gamma_4 + \gamma_5 n} (1 + p_t^f)^{\gamma_5 + \gamma_6 n}$$

where I do not assume that males and females labor supply are perfectly substitutable in the production function of home output.

Finally, assume that two-earners households cooperate so their preferences are defined by

$$U_{t,k}(c_t^m, c_t^f, d_t^m, d_t^f, p_t^f, Q_{t,k}, \theta_{t,k}) = \lambda_t * U_t^m(c_t^m, d_t^m, Q_t, \theta_{t,k}) + (1 - \lambda_t) * U_t^f(c_t^f, d_t^f, p_t^f, Q_t, \theta_{t,k})$$

where I assume that partners of the same type can marry and where  $\lambda_t$  is the pareto weight, determined endogenously as explained below.

### Incomes and Human Capital Accumulation

Family incomes are the sum of labor and non-labor incomes, discounting taxes and contributions. Labor incomes are driven by labor supply choices and the accumulated vector of skills of each household earner. I depart from a purely competitive market approach, so individuals receive job offers with stochastic arrival rates. This means that in every period with probability 1 individuals can work in self-employment activities if they want to, but job offers from the formal salaried and the informal salaried sectors arrive with certain probability which is determined exogenously to the model. I make these probabilities depend on age, education, gender, and the accumulated sector-specific experience. Thus

$$\Lambda_{a,t}^j(t^j, Ed^j) \text{ for } a = \{1, 2\}$$

Wages associated to a job offer are a function of the skill production function  $H_{a,t,k}^j$  and skill rental prices  $r_{a,t}$ . Skill functions vary by sector reflecting the existence of different production functions across sectors, where marginal productivities of each skill component, experience  $k_t^j$ , education  $Ed^j$  and abilities  $\mu_k$ , can differ in each sector  $a$ . Unobserved heterogeneity is incorporated into the skill function to capture self-selection into jobs based on sector-specific comparative advantage.

$$W_{a,t,k}^j = r_{a,t}^j H_{a,t,k}^j = r_{a,t}^j f(Ed^j, k_t^j, \mu_k)$$

Given a functional form for the skill function, the sector-specific log wage offer is defined by

$$\begin{aligned} \ln W_{a,t,k}^j &= \alpha_{a,0,k}^j \mu_k + \alpha_{a,1}^j E d^j + \alpha_{a,2}^j \ln(1 + k_t^j) + \epsilon_{a,t}^j \\ \epsilon_{a,t} &\sim N(0, \Sigma) \end{aligned}$$

where the idiosyncratic productivity shock  $\epsilon_{a,t}^j$  is assumed to be iid and serially uncorrelated.

Each household pays taxes according to the level of assets and labor incomes. Single-earner households face a tax function  $T^j(A_t, W_{1,t,k}^j, d_{1,t}^j, p_{1,t}^f)$ , where  $A_t^j$  is assets, while double-earners households pay taxes according to  $T(A_t, W_{1,t,k}^m, W_{1,t,k}^f, d_{1,t}^m, d_{1,t}^f, p_{1,t}^f)$ . Moreover, single and married households are eligible to observed transfer benefits like unemployment benefit or employment subsidies which can be conditional to marital status, the number of children at home and to the labor supply status. They are denoted by  $B^j(n_t, m_t, d_t^j)$ .

Therefore, single-earners household incomes are then defined by

$$y_t^j = \sum_{a=1}^{A-1} (1 - \phi d_{1,t}^j) W_{a,t,k}^j d_{a,t}^j p_{a,t}^j + B^j(n_t, d_t^j) - T_t^j$$

while double-earners household incomes are

$$y_t = \sum_{j=\{m,f\}} \left\{ \sum_{a=1}^{A-1} (1 - \phi d_{1,t}^j) W_{a,t,k}^j d_{a,t}^j p_{a,t}^j \right\} + B(n_t, m_t, d_t^m, d_t^f) - T_t$$

where pensions contributions paid over formal salaried wages are captured by  $\phi = 10\%$ .

### Wealth and Pensions

Family assets evolve according to standard budget constraints. However, as individuals can get married along the way and each of them carry her/his own accumulated assets, I assume that after marriage household assets are just the sum of the assets of each of the partners. As an example, the budget constraint for married households is

$$A_{t+1} = (1 + r_t)(A_t - c_t^m - c_t^f + y_t)$$

With regard to pensions, these are always gender-specific and evolve according to the monthly contributions paid every time individuals work as formal salaried and the earned return each period  $r^p$ , plus additional transfers  $\tau^j(n_t, m_t)$  reflecting the generosity of the pension system which may depend on the number of children



$n_t$  and the marital status  $m_t$ .

$$P_t^j = \phi W_{1,t,k}^j (d_{1,t}^j = 1) p_t^j (1 + r^p)^t + \tau^j(n_t, m_t)$$

### Children

The dynamics of fertility is determined by stochastic processes estimated outside of the model and incorporated exogenously. In particular, the probability of having children will be a function of the number of kids the year before, the marital status, the education and the age and labor status of the female in the previous period.

$$\pi_t^N = \pi^N(Ed_t^m, Ed_t^f, t^f, m_t, n_{t-1}, d_{a,t-1}^f)$$

## 5.2.4 Recursive Formulation

### The State Space

The state space for single households includes assets, the accumulated pension account, education, experience, the unobserved type, marital status, and the number of children.

$$\Omega_t^j = \{A_t^j, P_t^j, Ed_t^j, k_t^j, \mu_k, m_t^j, n_t^j\}$$

The state space for married households includes in addition the education, pension account and experience for both spouses, the matching quality and the bargaining power with which spouses enter in period  $t$ .

$$\Omega_t = \{A_t, P_t^m, P_t^f, Ed_t^m, Ed_t^f, k_t^m, k_t^f, \mu_k, n_t, m_t, \theta_{t,k}, \lambda_t\}$$

### Value Functions

Single and married individuals observe the shocks to wages and to utility costs before taking decisions. Individuals can enter period  $t$  single or married and they decide whether or not to continue in the same marital status. To do so, it's useful to defining the value function of staying single  $V_t^{j,S}(\Omega_t^j)$  by

$$\begin{aligned}
V_{t,k}^{j,S}(\Omega_t^j) &= \underset{s_t, d_{a,t}^j, p_{a,t}^j}{Max} \left\{ U_{t,k}^{j,S}(c_t^j, d_t^j, p_t^j, Q_{t,k}, \theta_{t,k}) + \beta \left[ EV_{t+1,k}^{j,S}(\Omega_{t+1}^j | \Omega_t) \right] \right\} \\
&\quad s.to \\
A_{t+1}^j &= (1 + r_t)(A_t^j - c_t^j + y_t^j) \\
P_{t+1}^j &= (1 + r_t^p)P_t^j + \phi W_{1,t,k}^j(d_{1,t}^j = 1)p_{1,t}^j + \tau^j(n_t, m_t) \\
Q_{t,k}^j &= Q(n_t, d_t^j, p_t^j, \mu_k) \\
A_t^j &\geq -B_t(y_t^j) \\
A_{T+1}^j &\geq 0
\end{aligned}$$

In a similar way, the value of getting divorced after marriage  $V_t^{j,D}(\Omega_t^j)$  is defined by

$$\begin{aligned}
V_{t,k}^{j,D}(\Omega_t^j) &= \underset{s_t, d_{a,t}^j, p_{a,t}^j}{Max} \left\{ U_{t,k}^{j,D}(c_t^j, d_t^j, p_t^j, Q_{t,k}, \theta_{t,k}) - D_t^j + \beta \left[ EV_{t+1,k}^{j,S}(\Omega_{t+1}^j | \Omega_t) \right] \right\} \\
&\quad s.to \\
A_{t+1}^j &= (1 + r_t)(\lambda_t A_t - c_t^j + y_t^j) \\
P_{t+1}^j &= (1 + r_t^p)P_t^j + \phi W_{1,t,k}^j(d_{1,t}^j = 1)p_{1,t}^j + \tau^j(n_t, m_t) \\
Q_{t,k}^j &= Q(n_t, d_t^j, p_t^j, \mu_k) \\
A_t^j &\geq -B_t(y_t^j) \\
A_{T+1}^j &\geq 0
\end{aligned}$$

where  $EV_{t+1,k}^{j,S}(\Omega_{t+1}^j | \Omega_t)$  is the expected value of entering period  $t + 1$  single. Two modifications follow from the previous case. First, divorced households must pay a utility cost of divorce  $D_t^j$ . And second, household assets are splitted between spouses according to the endogenous pareto weights.

To determine the individual value of staying married, I first define the problem of married households optimally deciding savings and labor supply denoted by  $z_t^* = \{s_t, d_{a,t}^m, d_{a,t}^f, p_{a,t}^m, p_{a,t}^f\}$  according to the following problem

$$\begin{aligned}
&\underset{s_t, d_{a,t}^m, d_{a,t}^f, p_{a,t}^m, p_{a,t}^f}{Max} \sum_{j \in \{m, f\}} (\lambda_t^j + \delta_t^j) \left[ U_{t,k}^{j,M}(c_t^j, d_t^j, p_t^j, Q_{t,k}, \theta_{t,k}) + \beta EV_{t+1}^j(\Omega_{t+1}^j | \Omega_t) \right] \\
&\quad s.to \\
A_{t+1} &= (1 + r_t)(A_t - c_t^m - c_t^f + y_t) \\
P_{t+1}^m &= (1 + r_t^p)P_t^m + \phi W_{1,t,k}^m(d_{1,t}^m = 1) + \tau^m(n_t, m_t) \\
P_{t+1}^f &= (1 + r_t^p)P_t^f + \phi W_{1,t,k}^f(d_{1,t}^f = 1)p_{1,t}^f + \tau^f(n_t, m_t) \\
Q_{t,k} &= Q_{t,k}(n_t, d_t^m, d_t^f, p_t^j, \mu_k) \\
A_t &\geq -B_t(y_t) \\
V_{t,k}^{j,D}(\Omega_t^j) &\leq U_{t,k}^{j,D}(c_t^j, d_{a,t}^j, p_{a,t}^j) + \beta \left[ EV_{t+1}^{j,M}(\Omega_{t+1}^j) \right] \\
\delta_{t+1}^j &= \delta_t^j + \Gamma_t^j
\end{aligned}$$

where the first four constraints are the standard constraints for a collective model, while the last two constraints define participation constraints to stay in marriage and will be explained below.

### Pareto Weights

Under this formulation,  $\lambda_t$  is defined as the pareto weight established optimally at the beginning of the marriage contract such that each of the spouses share equally the gains from marriage.

However, in further periods spouses can renegotiate this initial weight so that marriage continues and the participation constrain above holds for both spouses. Denoting  $\delta_t$  the acumulated deviation from the initial bargaining power, spouses enter period  $t$  with pareto weights  $\lambda_t + \delta_t$ .

To complete the description of the evolution of pareto weights, Ligon, Thomas, and Worrall (2002) show that in an optimal solution, Pareto weights can be increased by a social planner so that if the participation constraints binds for one of the spouses, the weight can be increased for the individual for that spouse so that it's still convenient for him/her to stay in marriage. In practice, defining  $\Gamma_t^j$  as the Lagrange Multiplier of the participation constraint above, the deviations from the initial weights evolve so that spouses remain in marriage as

$$\delta_{t+1}^j = \delta_t^j + \Gamma_t^j$$

and each spouse value function of marriage can be described by

$$V_t^{j,M}(\Omega_t^j) = U_{t,k}^{j,j}(c_t^{j*}, d_{a,t}^{j*}, p_{a,t}^{j*} | \delta_t^j + \Gamma_t^j) + \beta \left[ EV_{t+1}^{j,M}(\Omega_{t+1}^j | \delta_t^j + \Gamma_t^j) \right]$$

Instead, if pareto weights cannot be adjusted, then spouses leave marriage and their value function is  $V_{t,k}^{j,D}(\Omega_t^j)$ .

### Decisions

Individuals entering period  $t$  single can meet a potential partner with probability  $\pi_t^M$ , which is drawn from a distribution of assets, education and incomes for the two partners

$$\pi_t^M = \pi^M(A_t^m, A_t^f, Ed_t^m, Ed_t^f, y_t^m, y_t^f)$$

Each of these single-earner households decide whether to marry or not by comparing  $V_t^{j,M}(\Omega_t^j)$  and  $V_t^{j,S}(\Omega_t^j)$ , where the former is evaluated without including participation constraints.

Individuals entering period  $t$  married compare the value of staying married taking into consideration the participation constraints and the updated pareto weights, against the value of getting divorced, which accounts for utility costs of divorce and asset distribution upon separation.

### Expected Value functions

The individual Expected Value Functions after deciding marital status incorporate the probability of receiving job offers and can be expressed as

$$\begin{aligned} EV_{t,k}^j(\Omega_t^j) &= \Lambda_{1,t}^j \Lambda_{2,t}^j Emax \left[ V_{t+1,k}^j(\Omega_t^j, d_{1,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{2,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{3,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{4,t}^j = 1) \right] \\ &+ (1 - \Lambda_{1,t}^j) \Lambda_{2,t}^j Emax \left[ V_{t+1,k}^j(\Omega_t^j, d_{2,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{3,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{4,t}^j = 1) \right] \\ &+ \Lambda_{1,t}^j (1 - \Lambda_{2,t}^j) Emax \left[ V_{t+1,k}^j(\Omega_t^j, d_{1,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{3,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{4,t}^j = 1) \right] \\ &+ (1 - \Lambda_{1,t}^j) (1 - \Lambda_{2,t}^j) Emax \left[ V_{t+1,k}^j(\Omega_t^j, d_{3,t}^j = 1), V_{t+1,k}^j(\Omega_t^j, d_{4,t}^j = 1) \right] \end{aligned}$$

The two-earner Expected Value Function is defined similarly but accounting for sixteen combinations of family arrival jobs from the formal and the informal sector for the two spouses. and so on for all the possible combinations.

### 5.3 Final Comments

The state of the art in the understanding of the structure of labor markets in developing countries has for long been focused for on the study of partial incentives to labor participation like comparative advantage or the welfare system. Up to now, there is not such a study that attempts to put all the incentives together and that is able to provide answers from a life-cycle perspective and for different types of households. The model presented in this chapter undertakes that challenge and intends to apply the framework from the most advanced available models in labor economics

to serve the design of optimal labor and educational policies in the long-term. This is of course still work in progress, but I show it in here to present what will certainly be the central line of my research agenda over the next few years.

# References

- AAKVIK, A., K. G. SALVANES, AND K. VAAGE (2005): “Educational Attainment and Family Background,” *German Economic Review*, 6(3), 377–394.
- ABOUD, F. E. (2006): “Evaluation of an early childhood preschool program in rural Bangladesh,” *Early Childhood Research Quarterly*, 21(1), 46–60.
- ADDA, J., C. DUSTMANN, C. MEGHIR, AND J. ROBIN (2013): “Career Progression, Economic Downturns, and Skills,” *NBER Working Paper Series*.
- AGUIRREGABIRIA, V., AND P. MIRA (2010): “Dynamic discrete choice structural models: A survey,” *Journal of Econometrics*, 156(1), 38–67.
- ALTONJI, J. G., E. BLOM, AND C. MEGHIR (2012): “Heterogeneity in Human Capital Investments : High School Curriculum, College Major, and Careers,” *NBER Working Paper Series*.
- ALTONJI, J. G., F. HAYASHI, L. J. KOTLIKOFF, S. JOURNAL, AND N. DECEMBER (1997): “Parental Altruism and Inter Vivos Transfers : Theory and Evidence Parental Altruism and Inter Vivos Transfers : Theory and Evidence Fumio Hayashi,” *Journal of Political Economy*, 105(6), 1121–1166.
- AMARAL, P. S., AND E. QUINTIN (2006): “A competitive model of the informal sector,” *Journal of Monetary Economics*, 53(7), 1541–1553.
- ARBEX, M., A. F. GALVAO, AND F. GOMES (2010): “Heterogeneity in the Returns to Education and Informal Activities,” *INSPER Working Papers*, pp. 1–25.
- ARCIDIACONO, P., V. J. HOTZ, AND S. KANG (2012): “Modeling college major choices using elicited measures of expectations and counterfactuals,” *Journal of Econometrics*, 166(1), 3–16.
- ATTANASIO, O., AND K. KAUFMANN (2009): “Educational choices, subjective expectations, and credit constraints,” Discussion paper, National Bureau of Economic Research.

- ATTANASIO, O., C. MEGHIR, AND A. OTERO (2011a): “Formal Labor Market and Pension Wealth: Evaluating the 2008 Chilean Pension Reform,” Discussion paper, Mimeo. Department of Economics, UCL.
- (2011b): “Pensions, Work and Informality: The impact of the 2008 Chilean Pension Reform,” .
- BADEV, A., AND F. CUNHA (2012): “An Empirical Model of Parental Learning about the Technology of Skill Formation,” *Unpublished Manuscript*.
- BALDWIN, A. L. (1948): “Socialization and the parent-child relationship,” *Child Development*, pp. 127–136.
- BANDURA, A. (1995): *Self-efficacy in changing societies*. Cambridge University Press.
- BAUMRIND, D. (1966): “Effects of authoritative parental control on child behavior,” *Child Development*, pp. 887–997.
- (1968): “Authoritarian vs. authoritative parental control,” *Adolescence*.
- (1973): “The development of instrumental competence through socialization,” in *Minnesota symposium on child psychology*, vol. 7, pp. 3–46.
- (1989): “Rearing competent children.” .
- BECKER, G. S., AND N. TOMES (1979): “Becker-Tomes\_(JPoIE\_79).pdf,” *Journal of Political Economy*, 87(6).
- (1986): “Human capital and the rise and fall of families.,” *Journal of labor economics*, 4(3 Pt. 2), 1–47.
- BJÖRKLUND, A., AND K. G. SALVANES (2010): “Education and family background: Mechanisms and policies,” Discussion paper, Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit.
- BLACK, S. E., AND P. J. DEVEREUX (2011): “Recent Developments in Intergenerational Mobility,” in *Handbook of Labor Economics, Volume 4B*, 4(11), 1487–1541.
- BLAIR, C., AND R. P. RAZZA (2007): “Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten,” *Child development*, 78(2), 647–663.

- BLAIR, C., AND A. URSACHE (2011): “A bidirectional model of executive functions and self-regulation,” *Handbook of self-regulation: Research, theory, and applications*, 2, 300–320.
- BLUNDELL, B. R., L. PISTAFERRI, AND I. PRESTON (2008): “Consumption Inequality and Partial Insurance,” *American Economic Review*, 0214491(1998), 1887–1921.
- BLUNDELL, R., M. COSTA-DIAS, C. MEGHIR, AND J. SHAW (2013): “Female Labour Supply, Human Capital and Welfare Reform,” *NBER Working Paper Series*.
- BORNSTEIN, M. H. (2001): “Parenting: Science and practice,” *Parenting*, 1(1-2), 1–4.
- BORNSTEIN, M. H., O. M. HAYNES, H. AZUMA, C. GALPERÍN, S. MAITAL, M. OGINO, K. PAINTER, L. PASCUAL, M.-G. PÊCHEUX, C. RAHN, AND OTHERS (1998): “A cross-national study of self-evaluations and attributions in parenting: Argentina, Belgium, France, Israel, Italy, Japan, and the United States,” *Developmental Psychology*, 34(4), 662.
- BOSCH, M., E. GONI, AND W. MALONEY (2007): “The Determinants of Rising Informality in Brazil : Evidence from Gross Worker Flows,” *IZA Discussion Papers*, (2970).
- BOSCH, M., AND W. MALONEY (2007): “Comparative Analysis of Labor Market Dynamics Using Markov Processes : An Application to Informality,” *IZA Discussion Paper Series*, (3038).
- BRAVO, D., AND T. RAU (2012): “Effects of Large-scale Youth Employment Subsidies: Evidence from a Regression Discontinuity Design,” *Unpublished Manuscript*.
- BUTELMANN, A., AND F. GALLEGO (2000): “Household Saving in Chile: Microeconomic Evidence,” *Banco Central de Chile Documentos de Trabajo Central Bank of Chile Working Papers*.
- CAMERON, S. V., AND J. J. HECKMAN (2001): “The Dynamics of Educational Attainment for Black, Hispanic, and White Males,” *Journal of Political Economy*, 109(3), 455–499.




- CAMERON, S. V., J. J. HECKMAN, S. JOURNAL, AND N. APRIL (1998): "Life Cycle Schooling and Dynamic Selection Bias : Models and Evidence for Five Cohorts of American Males Life Cycle Schooling and Dynamic Selection Bias : Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106(2), 262–333.
- CARNEIRO, P., AND J. J. HECKMAN (2002): "THE EVIDENCE ON CREDIT CONSTRAINTS IN POST-SECONDARY SCHOOLING," *The Economic Journal*, 112(April 2001), 989–1018.
- (2003): "Human Capital Policy," in *Heckman, J. J. and Krueger, A. Inequality in America: What Role for Human Capital Policies*, MIT Press., (2000).
- CARNEIRO, P., K. V. LØ KEN, AND K. G. SALVANES (2011): "A Flying Start ? Maternity Leave Benefits and Long Run Outcomes of Children," *IZA Discussion Paper Series*, (5793).
- CAUCUTT, E., AND L. LOCHNER (2012): "Early and late human capital investments, borrowing constraints, and the family," *NBER Working Paper Series*.
- CHISLETT, G., AND D. J. KENNETT (2007): "The effects of the Nobody's Perfect program on parenting resourcefulness and competency," *Journal of Child and Family Studies*, 16(4), 473–482.
- CONTRERAS, D., L. D. MELLO, AND E. PUENTES (2008): "Tackling Business and Labour Informality in Chile," *OECD Economics Department Working Papers*, (607).
- CUNHA, F., I. ELO, AND J. CULHANE (2013): "Eliciting maternal expectations about the technology of cognitive skill formation," Discussion paper, National Bureau of Economic Research.
- CUNHA, F., AND J. HECKMAN (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.
- CUNHA, F., J. HECKMAN, AND S. SCHENNACH (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation.," *Econometrica*, 78(3), 883–931.
- CUNHA, F., AND J. J. HECKMAN (2006): "Formulating , Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *The Journal of Human Resources*, (December 2006).

- (2009): “The Economics and Psychology of Inequality and Human Development.,” *Journal of the European Economic Association*, 7(2), 320–364.
- DIAMOND, A. (1985): “Development of the ability to use recall to guide action, as indicated by infants’ performance on AB,” *Child development*, pp. 868–883.
- DUNCAN, G. J., W. J. YEUNG, J. BROOKS-GUNN, AND J. R. SMITH (1998): “HOW MUCH DOES CHILDHOOD POVERTY AFFECT THE LIFE CHANCES OF CHILDREN ?,” *American Sociological Review*, 63(3), 406–423.
- ENGLE, P. L., M. M. BLACK, J. R. BEHRMAN, M. DE MELLO, P. J. GERTLER, L. KAPIRIRI, R. MARTORELL, AND M. E. YOUNG (2007): “Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world,” *The Lancet*, 369(9557), 229–242.
- ENGLE, P. L., L. C. H. FERNALD, H. ALDERMAN, J. BEHRMAN, C. O’GARA, A. YOUSAFZAI, M. C. DE MELLO, M. HIDROBO, N. ULKUER, I. ERTEM, AND OTHERS (2011): “Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries,” *The Lancet*, 378(9799), 1339–1353.
- FERNALD, L. C. H., A. WEBER, E. GALASSO, AND L. RATSIFANDRIHAMANANA (2011): “Socioeconomic gradients and child development in a very low income population: evidence from Madagascar,” *Developmental science*, 14(4), 832–847.
- GALINDEV, R., AND D. LKHAGVASUREN (2010): “Discretization of highly persistent correlated AR(1) shocks,” *Journal of Economic Dynamics and Control*, 34(7), 1260–1276.
- GASPARINI, L., AND L. TORNAROLLI (2007): “Centro de Estudios Distributivos, Laborales y Sociales,” .
- GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): “Indirect Inference C. Gourieroux; A. Monfort; E. Renault *Journal of Applied Econometrics* , Vol. 8, Supplement: Special Issue on Econometric Inference Using Simulation Techniques. (Dec., 1993), pp. S85-S118.,” *Journal of Applied Econometrics*, 8.
- HAMADANI, J. D., F. TOFAIL, A. HILALY, S. N. HUDA, P. ENGLE, AND S. M. GRANTHAM-MCGREGOR (2010): “Use of family care indicators and their relationship with child development in Bangladesh,” *Journal of health, population, and nutrition*, 28(1), 23.

- HAYASHI, F. (2000): “Econometrics. 2000,” *Princeton University Press. Section, 1*, 60–69.
- HECKMAN, J., AND C. PAGÉS (2003): “Law and Employment: Lessons from Latin America and the Caribbean,” *NBER Working Paper Series*.
- HECKMAN, J., AND B. SINGER (1984): “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data,” *Econometrica*, 52(2), 271–320.
- HECKMAN, J. J. (2011): “Effective child development strategies,” *The Pre-K Debates: Current Controversies and Issues*, Paul H. Brookes Publishing Co., Baltimore.
- HECKMAN, J. J., AND R. PINTO (2013): “Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs,” Discussion paper, National Bureau of Economic Research.
- HECKMAN, J. J., AND E. VYTLACIL (2005): “Structural equations, treatment effects, and econometric policy evaluation,” *Econometrica*, 73(3), 669–738.
- HUMLUM, M. K. (2011): “Timing of Family Income , Borrowing Constraints and Child Achievement,” *Journal of Population Economics*, 24(3), 979–1004.
- IMAI, K., L. KEELE, D. TINGLEY, AND T. YAMAMOTO (2011): “Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies,” *American Political Science Review*, 105(04), 765–789.
- JENKINS, S. P., AND C. SCHLUTER (2002): “The Effect of Family Income During Childhood on Later-Life Attainment : Evidence from Germany,” *IZA Discussion Paper Series*, (604).
- JOUBERT, C. (2010): “Dynamic labor supply and saving incentives under a privatized pension system: evidence from Chile,” *Publicly accessible Penn Dissertations*, p. 112.
- JOUBERT, C., AND P. TODD (2011): “The Impact of Chile’s 2008 Pension Reform on Labor Force Participation, Pension Savings, and Gender Equity,” *Univ. of Pennsylvania, Working Paper*.

- JOVANOVIC, B. (1979): "Firm-specific capital and turnover," *The Journal of Political Economy*, pp. 1246–1260.
- KEANE, M., AND A. A. SMITH (2003): "Generalized Indirect Inference for Discrete Choice Models," .
- KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3).
- KENNAN, J. (2008): "Average Switching Costs in Dynamic Logit Models," (January).
- KILLINGSWORTH, M. R. (1987): "Heterogeneous Preferences, Compensating Wage Differentials, and Comparable Worth," *The Quarterly Journal of Economics*, 102(4), 727–742.
- LEVY, D. M., AND G. DUNCAN (2000): "Using sibling samples to assess the effect of childhood family income on completed schooling," Discussion paper, Northwestern University/University of Chicago Joint Center for Poverty Research.
- LEVY, S. (2010): *Levy, Santiago. Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico*. Brookings Institution Press, 2010.
- LIGON, E., J. P. THOMAS, AND T. WORRALL (2002): "Informal insurance arrangements with limited commitment: Theory and evidence from village economies," *The Review of Economic Studies*, 69(1), 209–244.
- MAGNAC, T. (1991): "Segmented or Competitive Labor Markets," *Econometrica*, 59(1), 165–187.
- MAGNAC, T., J.-M. ROBIN, AND M. VISSER (1999): "Analysing Incomplete Individual Employment Histories Using Indirect Inference," *Journal of Applied Econometrics*, 10.
- MALONEY, W. F. (1999): "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico," *The World Bank Economic Review*, 13(2), 275–302.
- MALONEY, W. F. (2004): "Informality Revisited," *World Development*, 32(7), 1159–1178.
- MANSKI, C. F. (2004): "Measuring expectations," *Econometrica*, 72(5), 1329–1376.

- MAYER, S. E. (1997): *What money can't buy: Family income and children's life chances*. Harvard University Press.
- MAZZOCCO, M., C. RUIZ, AND S. YAMAGUCHI (2014): "Labor Supply and Household Dynamics," *The American Economic Review*, 104(5), 354–359.
- MEGHIR, C., R. NARITA, AND J.-M. ROBIN (2012): "Wages and Informality in Developing Countries," *NBER Working Paper (Forthcoming American Economic Review)*.
- MEGHIR, C., AND S. RIVKIN (2010): "Econometric Methods for Research in Education," *National Bureau of Economic Research*.
- MOON, S. H. (2010): "Multi-dimensional human skill formation with multi-dimensional parental investment," *PhD diss., Department of Economics, University of Chicago*.
- MOSCOSO BOEDO, H., AND P. D ERASMO (2013): "Misallocation , Informality , and Human Capital : Understanding the Role of Institutions," *Virginia Economics Online Papers*, (301), 1–41.
- OHAN, J. L., D. W. LEUNG, AND C. JOHNSTON (2000): "The Parenting Sense of Competence scale: Evidence of a stable factor structure and validity.," *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 32(4), 251.
- PAGÉS, C., AND L. MADRIGAL (2008): "Is Informality a Good Measure of Job Quality ? Evidence from Job Satisfaction Data by Carmen Pagés," *Inter-American Development Bank*, (December).
- PAGÉS, C., AND M. STAMPINI (2007): "No Education , No Good Jobs ? Evidence on the Relationship between Education and Labor Market Segmentation," *IZA Discussion Paper Series*, (3187).
- PAULA, A. D., AND J. SCHEINKMAN (2007): "The Informal Sector : An Equilibrium Model and Some by The Informal Sector : An Equilibrium Model and Some Empirical Evidence from Brazil," *Penn Institute for Economic Research Working Paper*.
- PAXSON, C., AND N. SCHADY (2007): "Does money matter? The effects of cash transfers on child health and development in rural Ecuador," *World Bank, Washington, DC*.

- PERRY, G. (2007): *Informality: exit and and exclusion*. World Bank-free PDF, 2007.
- PISTAFERRI, L., C. MEGHIR, A. VOENA, AND H. LOW (2013): “Marriage, Social Insurance and Labor Supply,” in *2013 Meeting Papers*, no. 448. Society for Economic Dynamics.
- RAU, T., E. ROJAS, AND S. URZÚA (2013): “Loans for Higher Education: Does the Dream Come True?,” Discussion Paper 19138, National Bureau of Economic Research.
- RUPPERT, D., AND M. WAND (1994): “Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to The Annals of Statistics.  www.jstor.org,” *The Annual of Statistics*, 22(3), 1346–1370.
- SCHADY, N., J. R. BEHRMAN, M. C. ARAUJO, R. AZUERO, R. BERNAL, D. BRAVO, F. LOPEZ-BOO, K. MACOURS, D. MARSHALL, C. PAXSON, AND OTHERS (2014): “Wealth gradients in early childhood cognitive development in five Latin American countries,” *The Journal of Human Resources*.
- SHAPIRO, D., AND J. SLEMROD (1995): “Consumer Response to the Timing of Income : Evidence from a Change in Tax Withholding,” *American Economic Review*, 85(1), 274–283.
- SKRYPNEK, B. J., AND J. CHARCHUN (2009): *An evaluation of the Nobody’s Perfect parenting program*. Canadian Association of Family Resource Programs.
- SOLON, G. (1999): “Intergenerational Mobility in the Labor Market,” *In Handbook of Labor Economics, Volume 3. O. Ashenfelter and D. Card, editors.*, 3, pp.1761–1800.
- SOULELES, N. (1999): “The Response of Household Consumption to Income Tax Refunds,” *American Economic Review*, 89(4), 947–958.
- TAUCHEN, G. (1986): “Finite state markov-chain approximations to univariate and vector autoregressions,” *Economics Letters*, 20(2), 177–181.
- TETI, D. M., M. A. O’CONNELL, AND C. D. REINER (1996): “Parenting Sensitivity, Parental Depression and Child Health: The Mediation Role of Parental Self-Efficacy,” *Early Development and Parenting*, 5(4), 237–250.

- TODD, P. E., AND K. I. WOLPIN (2010): “Structural Estimation and Policy Evaluation in Developing Countries,” *Annual Review of Economics*, 2(1), 21–50.
- VOENA, A., AND D. BAYOK (2014): “Prenuptial Contracts, Labor Supply and Household Investments,” *Unpublished Manuscript*.
- WENTZEL, K. R. (1998): “Social relationships and motivation in middle school: The role of parents, teachers, and peers.,” *Journal of educational psychology*, 90(2), 202.
- WILLIS, R., AND S. ROSEN (1979): “Education and Self-selection,” *Journal of Political Economy*.
- YAMADA, G. (1996): “Urban Informal Employment and Self-Employment in Developing Countries : Theory and Evidence \*,” *Economic Development and Cultural Change*, 44(2), 289–314.

# Appendix A

## Additional figures Chapter 2

### Data descriptives

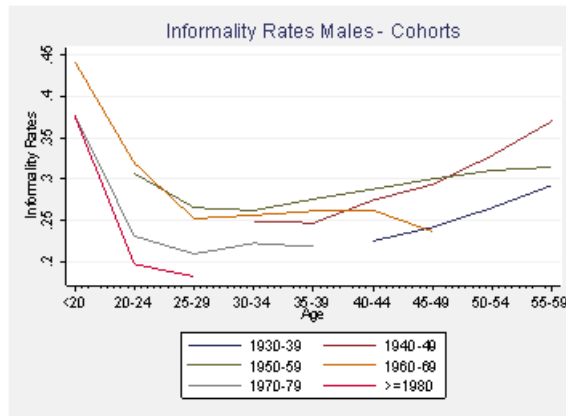


Figure A.1 Informality rates controlling by cohort effects.

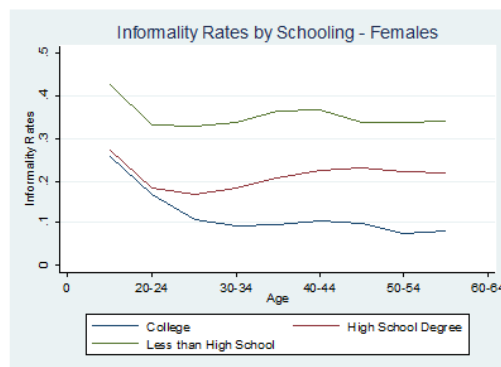


Figure A.2 Informality rates by schooling females



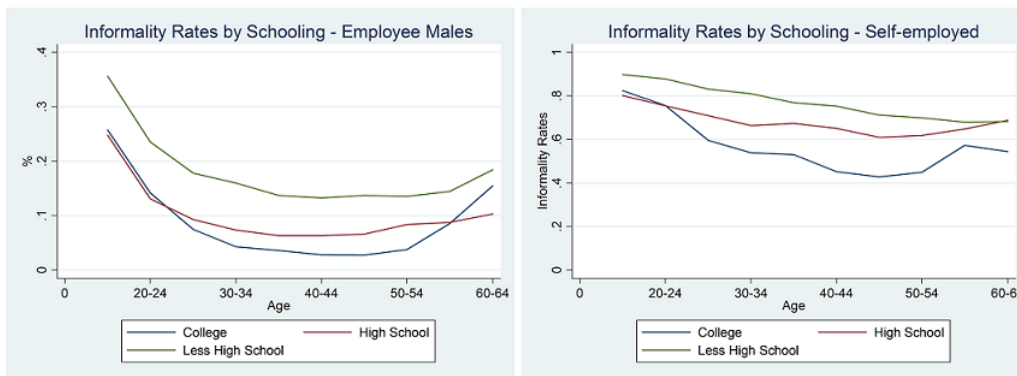


Figure A.3 Informality rates by schooling



Figure A.4: Wages by education, three sectors: Salaried Formal, Salaried Informal and Self-employed.

## Montecarlo Simulations Identification Exercise

Group	Parameters	$\theta_o$	$\theta_1$	$\hat{\theta}$	$\frac{\theta_1 - \theta_o}{\theta_o}$	$\frac{\hat{\theta} - \theta_o}{\theta_o}$
Psychic costs to schooling	$\gamma_{1,k=1}^{HS}$	1	1.200	1.021	20.0%	2.1%
	$\gamma_{1,k=1}^{Col}$	1	0.800	1.022	-20.0%	2.2%
	$\gamma_{1,k=2}^{HS}$	1.1	1.320	1.086	20.0%	-1.3%
	$\gamma_{1,k=2}^{Col}$	0.9	0.720	0.922	-20.0%	2.4%
Returns to Ability	$\alpha_{0,k=1}^F$	8.189	9.827	7.951	20.0%	-2.9%
	$\alpha_{0,k=1}^I$	8.007	6.406	8.089	-20.0%	1.0%
	$\alpha_{0,k=2}^F$	8.089	9.707	8.109	20.0%	0.2%
	$\alpha_{0,k=2}^I$	8.107	6.486	7.973	-20.0%	-1.7%
Returns to High School	$\alpha_1^F$	0.326	0.391	0.330	20.0%	1.1%
	$\alpha_1^I$	0.362	0.290	0.354	-20.0%	-2.3%
Returns to College	$\alpha_2^F$	1.307	1.568	1.294	20.0%	-1.0%
	$\alpha_2^I$	1.208	0.966	1.190	-20.0%	-1.5%
Returns to Experience same sector	$\alpha_3^F$	0.2	0.240	0.203	20.0%	1.5%
	$\alpha_3^I$	0.1	0.080	0.102	-20.0%	1.6%
Cross-sector returns to exp.	$\alpha_4^F$	-0.1	-0.120	-0.1016	20.0%	1.6%
	$\alpha_4^I$	0.1	0.080	0.101	-20.0%	0.8%
VCV wages	$\sigma_{FI}$	0.3	0.360	0.350	20.0%	16.8%
	$\sigma_F^2$	0.7	0.560	0.689	-20.0%	-1.6%
	$\sigma_I^2$	0.7	0.840	0.716	20.0%	2.3%
Tuition Costs Valuation	$\gamma_2^{HS}$	1	0.800	0.978	-20.0%	-2.2%
	$\gamma_2^{Col}$	1	1.200	1.019	20.0%	1.9%
Wage Valuation Informal	$\gamma_2^I$	0.7	0.560	0.687	-20.0%	-1.8%

Table A1: Montecarlo simulations estimation exercise

## The effects of extending the tax reduction to formal workers to the age of 40

Age	HS	Col
Total Sample	1.3%	0.5%
Type 1 (low)	1.5%	0.6%
Type 2 (high)	2.5%	-0.2%

Table A2: Extension of tax reduction to age 40 and schooling

Extension	Total Sample			Type 1			Type 2		
	LHS	HS	Col	LHS	HS	Col	LHS	HS	Col
14-17	-2.3%			-2.2%			-6.8%		
18-22	-3.1%	-3.1%		-3.1%	-2.9%		-9.0%	-2.1%	
23-26	-3.1%	-3.2%	-2.0%	-3.1%	-3.0%	-2.3%	-5.6%	-1.0%	-1.4%
27-30	-2.8%	-2.4%	-1.7%	-3.0%	-2.2%	-1.8%	-2.9%	-0.8%	-0.4%
31-35	-2.0%	-1.7%	-2.1%	-2.2%	-1.5%	-2.4%	-2.3%	-0.7%	-0.5%
36-40	-1.9%	-1.6%	-1.9%	-2.2%	-1.4%	-2.2%	-2.1%	-0.7%	-0.4%
41-45	-1.4%	-1.2%	-1.7%	-1.6%	-1.1%	-2.1%	-1.5%	-0.4%	-0.1%
46-50	-0.8%	-0.7%	-1.5%	-1.0%	-0.7%	-1.8%	-0.1%	-0.2%	0.0%
51-55	-0.5%	-0.3%	-1.2%	-0.7%	-0.2%	-1.5%	-0.1%	0.0%	0.0%
>55	-0.4%	-0.3%	-0.7%	-0.6%	-0.2%	-0.8%	-0.2%	-0.1%	0.1%

Table A3: Extension of tax reduction to age 40 and informality

# Appendix B

## Additional figures Chapter 3