Essays on early childhood interventions

Bloomfield Lescarbourá, J.

Publication date
2023

Document Version
Final published version

tinbergen institute

Essays on Early Childhood Interventions

Juanita Bloomfield Lescarbourá

Universiteit van Amsterdam

Essays on Early Childhood Interventions

Juanita Bloomfield Lescarbourá

ISBN: 978 90 361 0697 9

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 804 of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Essays on Early Childhood Interventions

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 25 januari 2023, te 13.00 uur

door Juanita Bloomfield Lescarbourá
geboren te Montevideo

Promotiecommissie

Promotor:	prof. dr. H. Oosterbeek	Universiteit van Amsterdam
Copromotor:	dr. A.I. Balsa	Universidad de Montevideo
Overige leden:	prof. dr. M. Lindeboom	Vrije Universiteit Amsterdam
	prof. dr. E.J.S. Plug	Universiteit van Amsterdam
	prof. dr. M.P. Pradhan	Universiteit van Amsterdam
	dr. G. Sorrenti	Universiteit van Amsterdam
	prof. dr. H.D. Webbink	Erasmus Universiteit Rotterdam

Faculteit Economie en Bedrijfskunde

Acknowledgements

I would like to start by expressing my sincere gratitude to my supervisors Hessel and Ana for their thoughtful advice and support during this process. Hessel, thank you for trusting that this experience of distance education was going to work out and for the freedom you gave me to pursue my research interests. I would also like to thank you for the time you have dedicated to give me detailed feedback and for always being available to discuss my work. I felt you were always "near" despite the geographic distance. Ana, I am very grateful for all the things that I have learnt from you. You have been invaluable both on an academic and a personal level. Feeling your support and trust has been very important to me during these years.

I gratefully acknowledge the confidence of Universidad de Montevideo and my professors in Uruguay. Alejandro, thank you for your support and for your positiveness at all times. Juan, I sincerely appreciate your help when applying to the PhD program and your feedback at different stages of my thesis. I would also like to thank the Uruguayan National Agency of Innovation and Research for their financial support in the form of a PhD grant.

I am grateful to the colleagues that have accompanied me during this process. To the Crianza Positiva Team for their companionship and commitment to the project. It has been a pleasure to work with you during all these years and I am very happy and excited that there is much more to come. Thank you to José María, for interesting discussions that benefited my learning experience.

A big thank you to my Tinbergen Institute's cohort for making the first two years of my PhD so much fun. I would like to especially thank Coen, Kim, Malin and my friends in Leiden, for so many good experiences during my days in the Netherlands. These are memories that will be treasured in my heart forever.

All my gratitude goes to my dear parents for their love, affection, and support in every step that I take. Mom and Dad, my most heartfelt thanks to you for *everything*. You are and you have always been there for me to listen, to help, to celebrate in the good times, to hug me and cheer me up in the low times, and much, much more: my gratitude to you is infinite.

Juanpi, thank you for listening about my research and for visiting me in the Netherlands. I really enjoyed this special time we had together.

Finally, I deeply thank my husband Carlos, my day-to-day companion in this journey. Thank you for pushing me forward, for encouraging me and supporting me. Looking back, I am very grateful for all the experiences I had the chance to live during my PhD, and I am happy that in all of these experiences you were by my side. To my daughters, Amalia, Clara and Antonia for bringing so much happiness to my life. Everyday is a joy because of you.

Contents

Chapter 1

Introduction

Over the last years there has been increasing evidence showing that early life conditions can have long-term consequences for human development and well-being. Investments during the zero to five period can be effective in shaping future outcomes and they are particularly relevant because of their potential to improve the situation of children that grow up in poverty. Policymakers around the world are implementing a wide variety of early childhood policies including cash transfer programs, the expansion of early childhood and preschool centers, and intervention programs with parents. That said, the question of how to best implement remediation in the zero to five period is not an easy one to answer. Causal evaluations of early childhood policies play an important role in the debate of where to invest and how to best invest.

This thesis includes three essays on the effectiveness of different early childhood interventions on parental investment, intergenerational transmission of poverty due to poor health at birth and educational attainment. Although each chapter is self-contained, they share at least two characteristics. First, they use data from Uruguay, a developing country in Latin America. In Uruguay, the existence of ability gaps in early childhood is a critical problem. There is evidence of large gaps in child development by socioeconomic level that persist throughout the school years. Second, all three essays focus on the estimation of causal impacts. I employ different identification strategies with the aim of giving a causal interpretation to my estimates.

The chapters in this thesis analyze programs that aim to help poor families by either providing parental tools to improve parenting practices and overcome behavioral biases, expanding access to local educational resources and providing cash to overcome financial constraints. Chapter 2 analyzes a program that was designed, implemented and evaluated by a group of economists and psychologists from the private sector using self-collected data while Chapters 3 and 4 evaluate the effectiveness of two public policies using administrative records. The specific contents of each chapter are described below.

Chapter 2 studies whether an e-messaging program rooted on behavioral economics insights is effective to increase parental investment and reinforce parental commitment. The messaging program is one of the components of *Crianza Positiva*, an intervention for parents of children aged 0-2 designed to promote good parenting practices. The intervention makes use of mobile messaging, a low-cost tool that can help foster parental engagement and contribute to habit formation. Treated families received messages three times a week for 24 weeks. The messages were designed based on behavioral economics theory and on the exploration of its predictions using baseline data. We conducted a preliminary analysis exploring evidence of behavioral biases in parenting behavior and found that present bias, inattention and cognitive fatigue, and negative identities were associated with the frequency of parental involvement. The contents of the messages aim to help parents reorient their attention towards positive parenting goals, simplify parental tasks, and reinforce positive identities.

The intervention was evaluated using a randomized controlled trial of the e-messaging program in 24 Child and Family Care Centers in Uruguay. The analysis uses self-collected data on the quantity and quality of parental investment, and on parental knowledge about positive parenting, self-efficacy, parental stress, and time preferences. Using this broad set of outcomes, the chapter sheds light on the relative strengths of different nudges by exploring the heterogeneity of parental responses when parents face different initial preferences, beliefs and constraints.

Chapter 3 estimates the effect of girls starting school earlier on health at birth of the next generation. I evaluate health at birth of the offspring of mothers that were exposed to a reform when they were 4 years old. The identification strategy uses a construction program of public preschool facilities implemented in Uruguay by the mid 90's that substantially increased the availability of preschool facilities. I exploit variation across regions and over time in the number of facilities built. The main database used in the analysis compiles information from vital statistics natality micro-data with a measure of availability of preschool places by region and year that I construct using school level data provided by the National Administration of Public Education.

Chapter 4 evaluates the long-term effects of receiving unconditional cash transfers since conception and up to the first five years of life on education outcomes. The focus is on the Uruguayan *Plan de Atención Nacional a la Emergencia Social* (PANES), a large unconditional cash transfer program that was implemented between 2005 and 2007 and that targeted the poorest 10 percent of households in the country. One special feature is that the amount of the transfer represented approximately 45% of the average household income among its target population. The analysis makes use of a dataset, specifically constructed for this project, that matches program administrative data to vital natality data and educational records 8 to 12 years after the beginning of the program. The identification strategy exploits a discontinuity that is generated from the eligibility rule of the program. The impact of the PANES program is separately estimated for children that were exposed to cash transfers since the in-utero period and children that were exposed later in life (but still in early childhood).

Finally Chapter 5 summarizes the main results and conclusions presented in the three chapters of this thesis.

Chapter 2

Using Behavioral Insights in Early Childhood Interventions: the Effects of *Crianza Positiva* E-messaging Program on Parental Investment[1]

2.1 Introduction

The importance of a nurturing environment for child development has been well established in the psychology, neuroscience, and economics literature. Research has also shown that this environment can be substantially enhanced by programs that expand and increase the quality of pre-school care, and by interventions that foster parental competences. Although the policy agendas in many countries have increased their focus on institutional early childcare, parenting programs are yet scarce, targeted at specific populations, and in many cases short-lived and too costly to apply. The challenge is to design cost-effective parenting interventions that can be scaled up to broad fractions of the population and are able to sustain parental behaviors over time.

In this chapter, we use random assignment to evaluate a text and audio messaging program aimed at reinforcing and sustaining positive parenting competences.[2] We assess the effect of the program on the quantity and the quality of parental investment reported by parents. The intervention is based on behavioral economics insights, in addition to early child development science. Its design recognizes that despite parents' good intentions, behavioral biases such as time inconsistency, cognitive fatigue and inattention, and negative identity, threaten the caregiver-child at-

[1] This chapter is based on Bloomfield et al. (2021)

[2] This random trial was registered in the American Economic Association RCT registry with ID number AEARCTR-0003585.

tachment, prevent parents from investing optimally, and ultimately affect the development of the child. By focusing on ways to overcome behavioral biases, behavioral economic interventions offer a set of promising tools to improve the environments in which children grow up and develop.

The messaging program that we evaluate is one of the components of *Crianza Positiva*, an intervention for parents of children aged 0-2 designed to promote positive parenting practices and competences. The program delivers voice and text messages to families for six months right after these families have completed an eight-week workshop at early childhood centers.[3] Previous evidence on the effects of the *Crianza Positiva* workshop show that while the workshop had significant and sustained benefits in terms of parenting knowledge and the quality of parental investment, it did not increase the frequency of parental involvement six months after the workshop ending (see Balsa et al. (2020)).[4] After workshop completion, we randomly assigned families to a treatment and a control group. Families in both groups participated in the 8-week workshop, but only those in the treatment group received text and audio messages after completing the workshop. Our analysis assesses the effects of the messaging intervention a year after parents initiate the workshop and three to five months after receiving the last message.

The covered topics in the messages include observing, interpreting, and responding sensitively to the child's signals, the importance of a safe and nurturing environment, the importance of speaking and reading to the child, the key role of free play, and the value of self-caring and of having a reflective parenting attitude. The messages, delivered three times a week, seek to help parents overcome behavioral biases by refocusing their attention towards positive parenting goals and the benefits of good parenting practices, by decomposing complex tasks into simpler ones, and by reinforcing positive parental identities.

The intervention makes use of mobile messaging, a low-cost tool that can help foster parental engagement and contribute to habit formation. Messages can provide continuous encouragement, support and reinforcement to parents over extended periods of time (York et al., 2019). In addition to their low cost, which makes them easily scalable, the use of mobile devices is widespread across large segments of the population, making the outreach of such programs nearly universal.[5]

[3]The workshop consisted of eight weekly sessions of 2.5 hours long. The design of the workshop does not explicitly focus on overcoming behavioral barriers in parenting.

[4]The first component of Crianza Positiva, the workshop, was not randomized, so in Balsa et al. (2020) we are unable to assess its impact using an experimental design. We evaluate treatment effects using matching techniques that compare families with similar observable characteristics at baseline that were exposed and not exposed to the workshop. We elaborate on this issue in the discussion at the end.

[5]Text and audio messages are a technological resource of high applicability in Uruguay where the use of cell phones is massive. The market penetration of cellphones, as measured by the quantity of unique connections over total population, is above 90% (D'almeida and Margot, 2018).

We find that the messaging component of *Crianza Positiva* increased both the quantity of parental investment, as measured by the frequency of parental involvement with the child, and the quality of parental investment, given by measures of parental outreach for social support and parental reflective capacity. The effects on the frequency of parental involvement are around 0.33 standard deviations and the effects on investment quality are around 0.31 standard deviations. We also explore some potential mechanisms behind the findings. Families initially exposed to high levels of negative shocks (our proxy for cognitive fatigue) or experiencing low parental self-efficacy (our proxy for negative beliefs) show larger responses to the messaging intervention. We do not find, on the other hand, differential responses by parental time discount rates.

Our study contributes to a growing literature on early-childhood interventions in several ways. First, our results support the use of text and voice messaging in combination with behavioral economics insights as a cost-effective tool to improve child-nurturing practices. We are upfront about the fact that our study cannot establish the stand-alone effects of the text messaging program. Because the messages were nested within a broader intervention that included a parenting workshop, our results could stem from the interaction between the learning that takes place in the workshop and the subsequent messaging program. Still, we believe that at a minimum, our results provide evidence about cost-effective ways of enhancing and sustaining over time the effects of face-to-face parenting programs, which have either been shown to fade out over time (as is the case of group-based parenting workshops) or have been quite expensive to maintain (as is the case of home visits). Second, there are a few recent papers on early-childhood parenting interventions that combine nudging with e-messaging for the United States. But we are unaware of other papers, apart from our own, showing similar evidence for the developing world. Third, and most important, we go beyond previous literature by exploring the heterogeneity of parental responses when parents face different initial preferences, beliefs and constraints. This helps us shed light on the relative strengths of different nudges. Our heterogeneity analysis suggests that nudges targeted at addressing cognitive fatigue, such as suggestions of simple actions, or nudges that boost parental self-esteem seem to be particularly effective in promoting parental investment. Finally, while most of the messaging programs we are aware of deal with parents of children that are at least in preschool, our program targets parents of children aged 0-2, an age when parental investment can have the largest returns.

The chapter is structured as follows. In Section 2.2, we provide a review of background literature, including the literature on socioeconomic gaps in early childhood investment and its relationship to child development, the recent findings of behavioral economics on parental decision-making, the use of technology in behavioral economics early childhood interventions, and the effectiveness of text-messaging programs. Section 2.3 provides a description of the *Crianza Positiva* program and of its text and audio messaging component. In Section 2.4 we describe the experimental design and evaluation instruments, assess compliance with randomization, attrition and balance, and present descriptive statistics of the data. Results are presented in Section 2.5, and we discuss and conclude in Section 2.6.

2.2 Background

In this section we present a literature review and a discussion of elements used in our text messages. First, we review the literature on the importance of early childhood for child development and the key role of parents during the early years. The findings from this literature are the main motivation for the intervention evaluated in this chapter. Second, we review the literature from the field of behavioral economics that explains how parental decision-making can be influenced by different types of biases. We provide a discussion of several elements that have been used to overcome these biases. The insights from these studies were used in the design of our text messages. Third, we review the literature of behavioral economics interventions that have used technology to increase parental involvement during early childhood. This chapter contributes to that strand of the literature. Finally, we provide a review of the literature on the effectiveness of text-messaging programs. We designed our text messages based on the findings from the latter studies.

2.2.1 Socioeconomic Gaps in Early Childhood and the Importance of Parental Investment

An extensive literature documents socioeconomic differences in parenting practices during early childhood. Families of low socioeconomic status spend less time with their children (Guryan et al., 2008; Schady et al., 2015b), show them fewer expressions of affection, and are more likely to physically punish them (Bradley et al., 2001). Low socioeconomic status families are also less likely to read to their children and to talk to them (Schady et al., 2015b; Bradley et al., 2001). In a seminal article, Hart and Risley (1995) show that, by the age of 4, children from low income families have heard about 30 million fewer words than other children.

Family environments in early life largely predict skill development. Heckman and co-authors underscore the importance of parental investment in the development of children's cognitive and non-cognitive skills, not only in the short-run, but also as a determinant of long-run social and economic success (Cunha et al., 2006). Children who grow up in a sensitive and stimulating environment tend to have better motor, social, emotional, numeracy and language skills development (National Institute of Child Health and Human Development Early Child Care Research Network, 2006; Bradley et al., 2001; Melhuish et al., 2008; Waldfogel and Washbrook, 2011; Schady et al., 2015a; Rubio-Codina et al., 2015). In Uruguay, socioeconomic ability gaps in early childhood were already documented three decades ago (Terra, 1988) and continue to be a critical problem (Uruguayan Nutrition, Child Development and Health Survey, 2018; López Bóo et al., 2019).

Formal schooling is unable to revert the socioeconomic gaps in ability that appear during the early years. Conventional indicators of school quality, such as the student-teacher ratio or teachers' salaries cannot compensate for early ability gaps (Heckman et al., 2005; Cunha and Heckman, 2008; Schady et al., 2015a; Tansini, 2008 and Llambí et al., 2009 for evidence for Uruguay). A systematic review by Almond and Currie, 2011 shows that the characteristics of the child and his family at school entry have as much predictive power as the years of education in explaining employment and wage outcomes. Furthermore, there is evidence that parental investment contributes to the intergenerational transmission of inequality (Holmlund et al., 2011).

Policies aimed at counteracting the effects of poverty on child development include cash transfer programs, the expansion and quality enhancement of early childhood centers and preschools, and intervention programs with parents (see Appendix 2.A for a review of the literature on parenting interventions). The evidence on the effectiveness of intervention programs with parents is vast (Nores and Barnett, 2010). However, these programs tend to have high costs (Leer et al., 2016) and difficulties to sustain effects over time (see overview by Bailey et al., 2017). The main challenge in the design of public policies aimed at enhancing parenting skills is to come up with interventions that are sufficiently low-cost to be scaled-up, but that can have sustainable effects over time. Behavioral economics and technology offer some promising tools to facilitate parental engagement at a low cost.

2.2.2 Behavioral Economics and Parental Decision-Making

To a large extent, parental interventions have been designed assuming that individuals act rationally (Gennetian et al., 2016). However, many decisions can be difficult to analyze and understand through the lens of the rational model. Recent findings at the intersection of psychology and economics -behavioral economics- are changing the way we understand how individuals make decisions and behave, offering new opportunities for the design of public policies (Rabin, 1998; Thaler and Sunstein, 2009; Congdon et al., 2011). Behavioral economics has grown rapidly due to its ability to explain sub-optimal outcomes and for its implications for public policies.

Public policies that incorporate behavioral economics insights have the potential to be highly cost-effective once they acknowledge that small changes in the way information is transmitted, or in the way choices are presented, can have large impacts on individual behavior.

Koch et al. (2015) and Lavecchia et al. (2016) identify key behavioral biases associated with education investment decisions and describe a set of tools that have proven effective in overcoming them. Appendix 2.B provides a detailed review of the literature from the field of behavioral economics and describes nudges to overcome biases in decision-making.

A key barrier to parental investment is present-bias (Thaler and Ganser, 2015). Families with large discount rates are less likely to invest in costly activities that provide benefits only in the future. The behavioral economics literature has proposed the use of commitments (Giné et al., 2010; Mayer et al., 2018), reminders (Cunha et al., 2017), and immediate incentives (Fryer Jr et al. (2015)) to overcome present bias.

Parental sub-optimal behavior can also arise from inattention and cognitive fatigue. The complexity of parenting may overwhelm and inhibit parents from making the right investment decisions, in particular when they are facing many stressors. To overcome the sub-optimality of decisions due to cognitive fatigue, Bryan et al. (2010) propose the use of reminder messages that make more salient the commitment to the desired objective, while Mayer et al. (2018), and York et al. (2019), propose designing solutions that facilitate parenting practices by decomposing complex tasks into simpler ones.

Negative identities may also act as a deterrent of optimal parental investment. Self-esteem and self-confidence are key factors to build intrinsic motivation. Positive feedback, motivational testimonies and peer support can help to promote positive identities (Lavecchia et al., 2016).

As many habitual behaviors, parenting is also subject to status quo bias or inertia. Adopting new parenting practices requires changing behaviors that are performed routinely, breaking with prior stereotypes, and defying the status quo (Samuelson and Zeckhauser, 1988). Establishing options by default can be an effective strategy to overcome this barrier (Madrian and Shea, 2001).

2.2.3 The Use of Technology in Behavioral Economics Interventions in Early Childhood

Interventions using technology to support parents are increasingly drawing attention due to their potential of expanding programs' reach at a low cost. Escueta et al. (2017) provide a review of the literature on interventions that use technology to support education decisions. The authors review five studies (described below) that experimentally evaluate programs that promote parental involvement in parents of preschool-aged children. They find positive effects in all the studies reviewed, which suggests that the use of technology holds great promise for early childhood interventions.

York et al. (2019) analyze the impact of "Ready for K!", an eight-month-long text-messaging intervention for parents of preschoolers. The program was implemented in San Francisco and aims to help parents support their children's academic development. The authors find that the intervention had a favorable impact on parental involvement at home and school, and improved child literacy skills. Doss et al. (2017) implemented the "Ready for K!" program in families with slightly older children and extended the program by adding a treatment arm in which the messages were personalized and differentiated according to the child's level of development. Although they find no impact of the original program "Ready for K!", they find that personalized messages increase the likelihood that parents read to their children by 50%.

In the same line of previous studies, Meuwissen et al. (2017) study the effects of the "Text2Learn" program: a 12-week text-messaging program for parents of low socioeconomic status in Minnesota, United States. The authors find that the program was successful in increasing the involvement of adults in the literary activities of their preschoolers.

Mayer et al. (2018) analyze the effect of a 6-week family intervention in which families were given a tablet with children's books. The treatment group received three weekly messages with: (i) reminders for the caregiver to read to the child, (ii) a tool to set weekly reading goals and give a feedback on the caregiver's progress, and (iii) messages of social gratification in the form of congratulations when the weekly objective was fulfilled. The authors find that parents in the treatment group used the tablets more than those in the control group and read more than twice as many books to their children. Hurwitz et al. (2015) study the impact of an intervention that consisted of daily messages to motivate parents to become more involved in didactic activities with their children and find positive impacts of this treatment.

While text-messages have shown promise to enhance parenting interventions for preschool- aged children, evidence on text-based interventions for new parents is just emerging. In their review of published research on technology-based interventions for parents of children aged 0-3, Hall and Bierman (2015) analyze studies with different technological approaches[6] and find only one study that evaluates a text-only messaging program. "Text4baby", consists of sending regular text-messages on topics such as prenatal behaviors, breastfeeding, developmental milestones and

[6] Including web-based platforms, discussion forums, mobile devices, and video conferencing.

infant care. Evans et al. (2012) show that females that receive the "Text4baby" messages feel more prepared for motherhood and that, among those with a high school education or higher, they also endorsed stronger attitudes against prenatal alcohol use. A more recent text-based approach for new mothers is "NurturePA", a mentorship program designed to support new mothers and promote healthy child development. Martin et al. (2018) provide a case study of NurturePA of the quantity and type of engagement a text-based mentorship intervention can support. The findings suggest that text-based mentoring is a promising strategy to engage and support new mothers.

2.2.4 The Effectiveness of Text-Messaging Programs

The effectiveness of text-message based programs is highly dependent on their design. Cortes et al. (2018) find that parenting programs based on text messages can provide excessive or insufficient information. A three-text-per-week approach that includes information, actionable advice and encouragement is more effective to improve parenting practices than approaches that include only one text per week or that include five. Fricke et al. (2018) analyze opt out of text messaging programs that aim to improve school readiness, and find that a high quantity of texts and more complex texts lead recipients to opt out more. Moreover, the authors find that programs that provide context and encouragement have lower opt out. In the context of a field experiment with charity, Damgaard and Gravert (2018) find that reminders via text messages and emails increase the intended behavior but also the avoidance behavior in terms of un-subscription from the mailing list.

2.3 *Crianza Positiva* and the Text and Audio Messaging Intervention

2.3.1 *Crianza Positiva*: a Multi-level Program Aimed at Strengthening Parental Competences

Crianza Positiva is a brief, preventive, highly protocolized and evidence-based intervention aimed at improving parenting practices and reinforcing child development. The program was originally designed to be implemented at "Children and Family Care Centers" (CAIF) of Uruguay. CAIF centers are publicly-funded, privately-managed[7] early childhood centers, whose purpose is "to guarantee the protection

[7] CAIF centers arise from an inter-sectoral alliance between Civil Society Organizations, the Uruguayan State, and Municipalities.

and promote the rights of children since their conception until the age of 3, prioritizing the access of those who come from families in poverty and/or social vulnerability".[8] We invited all CAIF centers across the country to participate in the *Crianza Positiva* intervention through a presentation at the CAIF National Committee. Because of limited resources, we chose to deliver the intervention to the first 24 CAIFs that expressed interest in participating. Families had no active role in enrollment decisions. They were automatically assigned to the program as long as they were attending "*Experiencias Oportunas*" (Timely Experiences), a weekly space at CAIF centers oriented to children aged 0-2 and their caregivers.

Crianza Positiva is a multilevel intervention, with varying treatment intensities according to the needs of each family. In the first level, all families participate in a workshop of eight weekly sessions, organized around four concepts: caring, stimulating, protective, and reflective competences. The workshop was designed to be delivered within the "*Experiencias Oportunas*" schedule.[9] The second level consists of a series of text and audio messages sent to workshop participants right after completing the workshop. This intervention seeks to help families incorporate the skills introduced in the workshop to their daily routines and nudge parents towards sustaining good parenting habits over time. This is the component that we evaluate in this chapter. At the third level, five parental counseling home visits are offered only to the most vulnerable families. The visits seek to deepen the development of the competences discussed in the workshop, accompanying the family in the management of sensitive observation, sensitive interpretation, and sensitive response to the child.

Crianza Positiva builds on the principles of positive parenting. Positive parenting encourages the creation of sensitive and structured environments at the family level, promotes the stimulation, support and recognition of children, and trains parents to be agents of change, competent, and able to positively influence their lives and the lives of their children. It builds on attachment theory, the theory of the mind and the ecological approach to parenthood. In relation to attachment theory (Main, 1991; Fonagy et al., 1991; Slade, 1999), the positive parenting principle seeks to encourage caregivers to be more skilled in the performance of their functions as facilitators of exploration and contributors of comfort and regulation of the child. Children who, during their first year of life, develop an insecurely attached relationship with their primary caregiver are at risk of deficits in socioemotional and cognitive development (Zeanah, 2000). Following Baron-Cohen's Theory

[8]http://www.plancaif.org.uy/plan-caif/que-es-plan-caif

[9]The workshop curriculum was based on the "Positive Parenting Scale Manual", developed by Gómez-Muzzio and Muñoz-Quinteros (2014), the "Nobody is Perfect" program in Canada, and the "Parents First program" (Goyette-Ewing et al., 2003), replicated in Finland under the name of "Families First".

of Mind (Baron-Cohen, 1997), the positive parenting principle seeks to strengthen the caregiver-child bond and good parenting practices through the stimulation of the reflective function of parents about the child's internal states and their role as caregivers. From an ecological perspective, positive parenting aims to help adults identify and rely on the ecological resources they have at their disposal (Bronfenbrenner, 1994). It also marks the commitment of community agents, who contribute from their role to the healthy development of parenthood.

2.3.2 The *Crianza Positiva* Text and Audio Messaging Component

The messaging component of *Crianza Positiva* consists of 72 messages sent to families three times a week over a period of 24 weeks. We chose to send three messages per week following the finding in Cortes et al. (2018) that the three-message-per-week approach is more effective than other approaches with fewer or more messages per week. The messages are sent right after families complete the workshop. Treated families receive the same messages both in text (via SMS) and in audio format (via WhatsApp).[10] Audio messages have exactly the same content as text messages, except that the latter are personalized with the name and the gender of the child. Messages are sent on Mondays, Tuesdays and Fridays at 6pm to all family members willing to receive them. Each family provides one or more mobile phone and messages are delivered to all family members willing to receive them.[11]

The 24 weeks of intervention are divided into 12 modules of two weeks. Each module refers to a different topic that was discussed in the *Crianza Positiva* workshop. Unlike most other messaging programs that target specific goals, such as reading and speaking to the child, our program covers a comprehensive range of parenting competences. These include sensitive observation and response, good treatment, safety at home, routines, speaking to the child, playing with the child, parental self-care, parental involvement, and parental reflection. Table 2.1 depicts the topics in the program by week of intervention.

The message structure was designed to address behavioral biases associated with low parental investment. Before designing the messages, we conducted a preliminary analysis exploring evidence of behavioral biases in parenting behavior. The analysis collected data through a self-administered questionnaire prior to families' partici-

[10]Messages were sent via text and WhatsApp to maximize the probability that the recipient received the message. Audio messages also reduce potential message failure in case the recipient was illiterate. Treated families received an opening message before the intervention and a closing message after the intervention. Four out of the 72 messages were sent in a video format to mobile phones that corresponded to females. These videos provided relaxing exercises. They were sent to females only because the images in the videos were related to females.

[11]The family could also opt out from the messaging intervention.

Table 2.1: Topics of messages by week of intervention and associated parental competence

Week	Topic	Parental competence
	Opening message	
1-2	Sensitive observation	Caring
3-4	Expressions of affection/sensitive approach	Caring
5-6	Safety and protection at home	Safety
7-8	The importance of routines	Safety
9-10	Self-caring for caring	Reflection/Self-care
11-12	Language: Speaking to the baby	Stimulation
13-14	Language: Reading	Stimulation
15-16	Free play	Stimulation
17-18	Relieving tensions	Reflection/Self-care
19-20	Learning how to calm oneself	Reflection/Self-care
21-22	Parental involvement	All competences
23-24	Reflection about parenting	Reflection
	Closing message	

pation in the *Crianza Positiva* workshop. Results from this analysis are presented in Appendix Table 2.i.[12] We found that the frequency of parental involvement in stimulating activities with the child correlated negatively with the parents' time discount rate, suggesting that present-oriented caregivers place a higher weight on the current costs of parental investment relative to future benefits.[13] Parental investment in stimulating activities decreased also with parental stress, suggesting that inattention and cognitive fatigue could be behind sub-optimal parenting decisions (Cooklin et al., 2012). We identified a similar negative relationship between stress and parental involvement in physical games. Lastly, parental involvement in physical games and social activities correlated positively with parents' sense of competence, suggesting that parental identity and sense of self-efficacy play a role at some level on parental investment decisions.

Based on behavioral economics theory and on the exploration of its predictions using the baseline data, we built messages around the following theory of change: (i) reminding parents about the benefits of engaging in positive parenting behavior will make these benefits more salient and tangible, in particular for present-biased parents, and improve investment through activation and recall of prior knowledge; (ii) providing parents with suggestions of simple and concrete activities will help address inattention and decision fatigue by decomposing the complexity of parenting

[12]The scales used to assess parental involvement, parenting stress, discount-rate, and other measures described in Appendix Table 2.i are discussed in more detail in section 2.4.1.

[13]A negative association between parental investment and the discount rate is a necessary but not sufficient condition for time inconsistency.

into simpler tasks; (iii) providing parents with self-care suggestions and techniques will help address inattention and decision fatigue by improving parental self- control and emotional regulation; (iv) encouraging parents to continue trying, telling them that they have the resources their child needs, that they are not alone in facing difficulties and making mistakes when it comes to raising children, and that it is good to rely on others will help them overcome negative identities and strengthen parental self-efficacy.

The structure of the messages followed a 2-weeks pattern. During the first week, the Monday message sent information on the importance of a certain parenting competence. On Tuesday, parents received a suggestion to engage in a specific activity with the child in relation to that competence. On Friday, parents were invited to reflect on their performance during the week and on their personal feelings regarding the task proposed, and usually received an encouragement message. The Monday message on the second week reinforced parental self- efficacy; on Tuesday a new task on the same topic was suggested; and the last message on Friday reinforced the importance of the parental competence discussed in the 2-week module and motivation to continue practicing in the future (see Table 2.2 for examples of messages).

The messages included a few other components that aimed to strengthen personal commitment and to provide parents with other sources of information and ideas. First, in the last day of the *Crianza Positiva* workshop, participants were asked to choose three behaviors that they could commit to practice in the following months and that they would like to be reminded of in the future. The options involved behaviors related to the topics covered in the messages. We used these selections to send each treated family a reminder of their commitments on the last day of the bi-weekly module.

In addition, treated families were provided a username and password via SMS to access "*Radio Butiá*", a Uruguayan web server that hosts Latin American songs and stories online. We also directed families via SMS to access a virtual platform in Facebook to find additional information about specific topics that were mentioned in the messages. We updated this information every week. This page was mostly informative and did not encourage interaction.

2.4 Impact Evaluation: Methodology

2.4.1 Experimental Design

Sample and Randomization

The *Crianza Positiva* workshop took place between September and November 2017 at 24 CAIF early childhood centers. After workshop completion, we conducted a

Table 2.2: Examples of messages by type of behavioral bias

Behavioral bias	Type of message to address this barrier	Example
Present bias and time inconsistency	• Messages that underscore the benefits of parenting practices • Reminders about the importance of parenting practices	*Children's brains are like sponges, they absorb everything: the sounds, the pitches of voices, the language they listen to. The more words your child listens to at this stage, the more [she] will develop [her] language. It is therefore very important that you speak to [child's name], this will impact heavily on [her] ability to learn.*
Complexity of parental role, inattention, decision fatigue	• Messages that decompose complex parental tasks into simple ones • Messages that relieve stress through breathing and relaxation techniques	*Talk to [child's name] while you are washing [him] or changing [his] diapers. Look [him] in the eye when you speak to [him]. When [child's name] tries to respond, don't interrupt [him] and do not get distracted. [Child's name] needs to know you are listening.*
Negative identity	• Messages that strengthen parental self-efficacy and empowerment • Messages showing that feeling stressed out or underconfident is normal, and that parents do overcome it	*There is no one that wants as much for [child's name] as you do. Think about one or two moments in the past days in which you felt you really contributed towards [her] wellbeing. Trust yourself and continue seeking more of these moments during the week.*
Status quo bias	• Messages suggesting concrete and simple activities • Messages reminding benefits of parental involvement	*The more you speak to [child's name], the better will [his] language develop and the more [he] will learn. Today and in the following days, remember and repeat this thought: "I take advantage of all the moments with [child's name] to speak to [him]."*

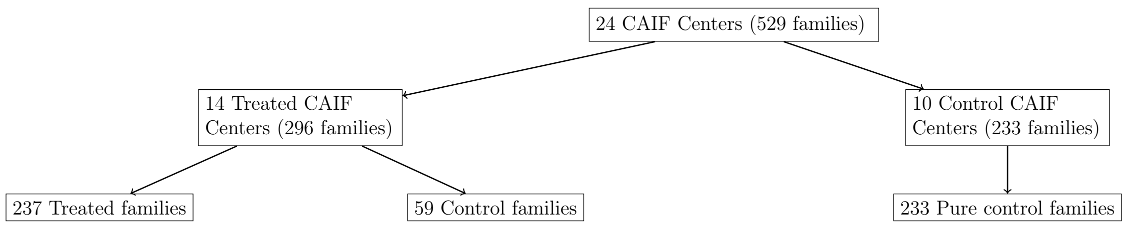
randomized controlled trial (RCT) to assess the effects of the *Crianza Positiva* text and audio messages component.[14] We opted for a two-stage randomization design that would enable us to assess the degree of spillovers from treated families to control families (Baird et al., 2018) in addition to the main effect of the messaging intervention. 529 families from the 24 CAIF centers were randomized to treatment in two steps (Figure 2.1 illustrates the randomization process). First, we stratified early childhood centers by average maternal education[15] and within each strata randomly

[14]The project was approved by the Institutional Review Board at Universidad Católica del Uruguay, resolution A 22-08-17.

[15]As a proxy for socioeconomic status, we used mother's average years of completed education of children that attended the early childhood centers. We constructed two strata according to whether the average of years of education was equal to or above nine years. One stratum had six CAIF centers and the other had eight. We obtained this information from baseline data which was collected before the messaging intervention.

allocated 60% of centers to a treatment arm and 40% of centers to a control arm. As a result of the first step of the randomization, 14 early childhood centers (296 families) were assigned to treatment and 10 centers were assigned to a pure control (none of the 233 families in these centers were treated). Secondly, within each center in the treatment arm, we randomized families into receiving or not receiving messages. In centers that were treated, 80% of families were selected to receive messages.[16] From this second stage, 237 families were randomly assigned to receive messages and 59 were assigned to control.

Figure 2.1: Randomization



Messaging Intervention Timeline and Implementation

The messaging intervention took place between January 5, 2018 and June 27, 2018. The school year in Uruguay runs from March to December, so one third of the program was delivered during summer holidays. We asked families to provide us with at least two mobile numbers where they could be reached. When no information was available, we asked the CAIF center to provide us with the numbers. We delivered the intervention to all the contact numbers we had for each family (mothers, fathers and other caregivers). This allowed us to maximize the chances of reaching the family and to increase males' participation (most of the participants at the *Crianza Positiva* workshop were women). In total, we had 373 mobile phone numbers corresponding to 237 treated families.

We sent all treated families a welcome message (both via SMS and WhatsApp) on January 5, 2018 and a closing message on June 27, 2018 that thanked them for participating in the program. Control families received a single SMS message on January 26, 2018 thanking them for participating in the *Crianza Positiva* workshop.

Text messages were delivered through a platform that enabled us to send the same message to all families at once. Audio messages were sent via a WhatsApp

[16]The proportion of centers assigned to pure control and the saturation of the treatment within treatment centers were selected in order to maximize power. Power calculations were performed using the Matlab code provided in the Supplemental Appendix of Baird et al. (2018).

broadcast list. Due to the specific characteristics of broadcast lists in WhatsApp, these messages could only be delivered to recipients who had saved the *Crianza Positiva* phone number in their contact list.[17]

Compliance

SMS messages were sent as planned, but we could not control whether they were delivered or received. Messages could fail because of incorrect or unavailable mobile number, busy line, or no credit. We re-contacted all treated families by the end of January 2018 and randomly selected a sample of treated families in March 2018 to verify whether they were getting the messages. By January, we detected that 17% families were not receiving any SMS messages (40 out of 237 treated). Whenever we could identify that messages were failing, we asked the early childhood center to update families' mobile numbers.If we exclude from the analysis those families that were not receiving either SMS nor WhatsApp and that started receiving messages after we updated their contact numbers, our results remain qualitatively the same. This suggests that if the intervention would not have had involved the additional steps taken to ensure that the intervention was reaching families, we should not expect smaller impacts on average. For 23% of these families, we could update at least one line, which reduced our SMS failure rate to 13% of families. Regarding WhatsApp messages, we found that in 87% of families, at least one member of the family received the messages. We also found that families that received WhatsApp messages read on average 69% of the them. Overall, we found that 11% of families (27 out of 237) did not receive SMS or opened WhatsApp messages. In Appendix Table 2.ii we provide a first stage regression where the outcome variable is an indicator for whether the family actually received messages and the independent variable is the assignment to treatment (or intention to treat) status.

We had an additional source of failure with messages sent to cellphones belonging to one of the carriers[18] between 30 January and 20 March 2018 (36% of our sample). The carrier labeled our messages as "spam" and the messages were not delivered. However, 64% of these cellphones kept receiving WhatsApp messages during this period. Furthermore, the randomization was balanced across different carriers.

Regarding the Facebook component of the intervention, we found that 83 families (35%) signed up to the *Crianza Positiva* Facebook informative web. We were unable to assess which families downloaded the "Radio Butiá" stories and songs.

[17]In our welcome SMS message to the program, we included our cellphone number and asked recipients to save our contact phone in order to keep receiving messages through this channel.

[18]In Uruguay there are three carriers: Ancel, Claro, and Movistar. The problem appeared with cell phones carried by Movistar.

Evaluation Scales

We collected data on the quantity and quality of parental investment, and on parental knowledge about positive parenting, self-efficacy, parental stress, and time preferences through a self-administered questionnaire. We collected these outcomes at baseline (before the messaging intervention) and in a follow-up survey administered at least two months after the messaging intervention had ended (between August and November 2018). Figure 2.i in the Appendix shows the timeline of the entire project indicating when each treatment took place and when which data were collected.

The questionnaire had to be completed by the parent or another caregiver in the presence of the interviewer either at home or at the early childhood center (depending on family preferences) and took approximately 40 minutes to complete.[19] A general description of the questions included in the questionnaire is provided below. For further information on the construction of each instrument please refer to Appendix Table 2.iii.

A sociodemographic section contained questions that covered demographic characteristics of the child and the respondent (usually the child's mother), the relationship between the respondent and the child, and household characteristics, such as household composition, maternal and paternal education, maternal and paternal employment, household assets, indicators of unsatisfied basic needs in the household, and cash transfers recipient status.

To evaluate the quantity of parental investment we considered the following dimensions. First, we inquired about the frequency of parental involvement in physical, social, caring, and didactic activities with the child. These questions were taken from Cabrera et al. (2004) and were previously used in the evaluation of Early Head Start in the United States. Second, we included items from the Father's Involvement subscale of the Etxadi-Gangoiti Scale (Arranz Freijo et al., 2012), which gathers information about the participation of the father in the nurturing of the child, in household tasks, and the quality of his interaction with the child. Third, we asked parents about material resources at home, such as availability of books and different types of toys (role playing toys, push or pull toys, musical instruments, etc), with which we built a material resources index.

To assess investment quality, we administered a subset of items from the Positive Parenting Scale (E2P), by Gómez-Muzzio and Muñoz-Quinteros (2014). The manual classifies the parent-child relationship in four groups: (i) nurturing and attachment, (ii) safety and protection, (iii) stimulation, and (iv) parental reflective capacity. We also added a set of items from UNICEF MICS6 questionnaire for families of children

[19]Parents had to sign an informed consent prior to answering the questionnaire.

under five, inquiring about the disciplinary methods that parents used with their child in the month before the survey.

To elicit time preferences, we administered the Kirby, Petry, and Bickel's (Kirby et al., 1999) Monetary Choice Questionnaire (MCQ). The instrument identifies a time discount rate for each individual that ranges from 0 to 0.249. A higher value indicates a higher preference for the present. The survey also included Abidin's (Abidin, 1995) Parenting Stress Index (Short Form) (PSI/SF), an instrument for parents with children between the ages of one month and 12 years old, designed to assess stress experienced when exercising parenting activities. The scale is divided in three subscales (i) "Parental discomfort" which identifies the discomfort that parents experience when performing parenting duties and is derived directly from personal factors that are related to parenting (tensions or conflicts), (ii) "Dysfunctional Interaction between parent and child" which assesses whether children meet parents' expectations and the degree of satisfaction that parents have with the child, and (iii) "Difficult Child" which identifies whether the caregiver considers the child-rearing tasks easy or difficult. The scale includes, in addition, a set of questions about stressful events that the household faced in the last 12 months. The latter answers are not considered in the overall score of the PSI/SF but are used in the analysis as exogenous sources of stress.

Parents' perceptions about their competence as parents were gathered with the Johnston and Mash (Johnston and Mash, 1989) version of the Parental Sense of Competence Scale (PSOC). We constructed two subscales of the PSOC suggested by Menéndez et al. (2011), one related to "effectiveness" and the other one capturing "controllability" of the parenting role. The former captures whether the adult feels capable and competent to act effectively as a parent. Controllability is determined by the degree to which parents feel responsible for education situations and consequences.

Finally, we assessed parental knowledge about positive parenting by including 13 True/False items.

2.4.2 Attrition

Out of the 237 families randomized to receive messages (ITT=1), 72% responded to the follow-up questionnaire (171 families). The response rate for the 292 families randomized to the control arm was 78% (see Table 2.3). This difference in attrition between treated and control subjects is not statistically significant at usual levels. However, if we analyze this for each outcome and take into account missing responses in the questionnaire, some of the outcomes show larger differences (with up to a nine percentage points difference). In the next section, we assess balance in covariates across treated and control families that responded to the questionnaire and check also for balance considering the missing values in different subsets of observations.

Table 2.3: Attrition by ITT

	Response to follow-up
Coefficient	-0.056
	(0.038)
Constant	0.777***
	(0.024)
N	529

Note: Table shows results from a regression of a dummy variable that takes the value of 1 if the family responded to the questionnaire and 0 if not on the ITT variable and a constant. The coefficient reported can be interpreted as the difference in response rate between ITT=1 and ITT=0. The constant can be interpreted as the response rate in the control arm. N=Number of observations. Standard errors in parentheses. * $p<0.1$, ** $p<.05$, *** $p<.01$

2.4.3 Descriptive Statistics and Balance

We begin by describing evaluated families and children according to a set of sociodemographic indicators reported by the family between August and December 2017 (note that the first message was sent in January 2018). We then use this data to check for randomization balance after accounting for attrition.

Table 2.4 presents descriptive statistics of the respondents, the children and the households at baseline, by ITT status. Mothers are around 29 years old and children are two years old on average at the time of initiation of the messaging intervention. Eight percent of the children were born prematurely. Three out of four children live with their biological father and mother, and one out of four are still being breastfed by the time the intervention begins. On average, there is one other child in the household aside from the evaluated child and 0.2 other adults aside from the child's parents. One out of three households faces material housing problems (problems in walls or floors, overcrowding or lack of a space to cook); only 2% lack running water, 3% lack sanitation, and 21% have no access to at least one basic comfort asset, including heating, a fridge, and a water heater. We construct an asset index including 18 household and family assets and utilities.[20] The index ranges from 0 to 0.77 with higher values indicating higher availability of assets. The average value of the asset index for families participating in the study is 0.25 with a standard deviation of 0.14. Sixty six percent of families are recipients of government cash transfers. Eighty percent of families attended six or more sessions of the eight-session workshop. Mothers are by far the main respondents to the questionnaire (93%), followed by fathers, grandmothers and other caregivers. The average time in months to survey response, since January 1, 2018, was 8 months.[21] The child under study is the mother's only child in 38% of the cases. Almost 30% of mothers are high school graduates; one out of three completed middle school but not high school, and the rest did not complete middle school.[22] Fathers work full time in 79% of the cases. Nearly three out of five families report having experienced a negative shock in the past year (a death, a divorce, unemployment, money problems, problems with the law or with drugs in the family); the average number of problems is 1.34. On average, 38% of families are classified as having a high discount rate, meaning that the discount rate of the respondent is higher than 0.1 (discount rates range from 0 to 0.249). Lastly, around half of respondents have a low parental sense of competence at baseline, which implies scoring four or less in the subscale of efficacy of the Parental Sense of Competence Scale.

[20]These include oven, refrigerator, water heater, TV, DVD, subscription to cable TV, laundry washer, laundry heater, dishwasher, microwave, air conditioner, government awarded laptop, other laptops, access to Wi-fi, household phone line, motorcycles, and cars.

[21]Participants responded to our follow-up survey between August and November 2018.

[22]In Table 2.iv we show a comparison between the descriptive statistics of the main sample and the general population using data from the Uruguayan National Survey of Nutrition, Child development and Health of 2018. In 45% of Uruguayan families with children aged 0-36 months the mother has completed middle school education, and in 27% the mother is a high school graduate. If we take maternal education as a proxy for socioeconomic status, the latter statistics show that families that attend CAIF are relatively disadvantaged in comparison to the general population.

Table 2.4: Descriptive statistics and balance in covariates across treatment arms

	Treatment			Control			Full sample balance		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Difference	Diff. s.e.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's age	168	30.292	6.626	225	28.311	6.552	393	1.981***	(0.672)
Female child	170	0.482	0.501	227	0.498	0.501	397	-0.015	(0.051)
Child's age (in months)	167	23.343	7.014	226	24.419	6.115	393	-1.076	(0.678)
Premature child	170	0.065	0.247	225	0.089	0.285	395	-0.024	(0.027)
Intact family	167	0.719	0.451	224	0.777	0.417	391	-0.058	(0.045)
Child still being breastfed	170	0.253	0.436	226	0.248	0.433	396	0.005	(0.044)
Number of other children in household	159	1.138	1.285	216	0.903	1.151	375	0.236*	(0.129)
Other adults in household	158	0.158	0.366	215	0.228	0.420	373	-0.070*	(0.041)
Housing material problems	158	0.291	0.456	218	0.303	0.461	376	-0.012	(0.048)
No running water	168	0.018	0.133	226	0.022	0.147	394	-0.004	(0.014)
No sanitation	168	0.024	0.153	226	0.031	0.174	394	-0.007	(0.017)
No basic comfort goods	159	0.226	0.420	223	0.202	0.402	382	0.025	(0.043)
Asset index	150	0.241	0.140	222	0.256	0.139	372	-0.015	(0.015)
Beneficiary of cash transfers	170	0.700	0.460	227	0.634	0.483	397	0.066	(0.048)
Attendance to $\geq$ 6 workshop sessions	169	0.793	0.406	223	0.820	0.385	392	-0.028	(0.040)
Mother responds questionnaire	170	0.935	0.247	227	0.916	0.278	397	0.019	(0.026)
Average time (in months) to survey response	166	7.880	1.119	226	7.882	1.079	392	-0.000	(0.113)
Only child	163	0.337	0.474	220	0.414	0.494	383	-0.076	(0.050)
Mother completed middle school	167	0.317	0.467	226	0.332	0.472	393	-0.014	(0.048)
Mother completed high school	167	0.281	0.451	226	0.305	0.462	393	-0.024	(0.046)
Mother works full time	165	0.388	0.489	217	0.401	0.491	382	-0.013	(0.051)
Mother works part time	165	0.194	0.397	217	0.194	0.396	382	0.000	(0.041)
Father works full time	151	0.808	0.395	215	0.767	0.423	366	0.041	(0.043)
Family had $\geq$ 2 negative shocks in 12 months	74	0.311	0.466	104	0.327	0.471	178	-0.016	(0.071)
High discount rate	138	0.399	0.491	194	0.366	0.483	332	0.033	(0.054)
Low sense of parental competence	74	0.527	0.503	94	0.447	0.500	168	0.080	(0.078)

Note: N=Number of observations. Std. Dev.=Standard Deviation. Column (8) reports the difference in means between Treatment and Control arms. Column (9) shows the standard errors associated with the coefficients in Column (8). * $p<0.1$, ** $p<.05$, *** $p<.01$

The last two columns in Table 2.4 show the differences in covariates at baseline between treatment and control subjects that responded to the follow-up questionnaire. Out of 26 covariates analyzed, only mother's age is statistically different at 1% across treatment and control subjects (mothers in the treatment group are almost two years older than mothers in the control group). Two other variables are statistically different at 10%, the number of other children in the household (which is larger for treated families), and the number of other adults in the household (which is smaller for treated families). We adjust for these three covariates in the regression analysis.

Table 2.5 shows descriptive statistics of outcomes assessed at follow-up. Groups of outcomes are identified by a heading in bold. The parental investment subscales indicate the frequency with which parents engage in different parenting activities, with one indicating "Never" and six "All or most days of the week".[23] The Parental Time Investment Index is a summation of the physical games, didactic and social activities scales: it averages 13 and has a maximum of 18. Father's involvement in childrearing is a continuous index from zero to one with an average of 0.77. 84% of families have at least five children books in the household.

The Positive Parenting Scale (E2P) shows values above four (on a maximum scale of five) for attachment, and values above 3.5 for routines, social support and parental reflection.[24] Parents rate higher levels of stress when asked about personal discomfort and child characteristics, than when assessing the interaction with the child. The average sense of parental competence is four, on a scale from one to six. The percentage of families reporting the use of some type of violence as a "disciplinary" approach was 39%.

Parenting knowledge is the summation of different true-false items indicating knowledge of positive parenting competences. On average, parents had 11 out of 13 questions right. Finally, the time discount rate (the rate at which parents discount the future) averages 0.08 in the sample, with a minimum of 0 and a maximum of 0.249.

[23]Because almost all parents report taking care of the child all or most days of the week, we exclude this outcome from the analysis.

[24]Note that the Positive Parenting Scale is not the original scale in Gómez-Muzzio and Muñoz-Quinteros (2014), but a subscale constructed on a subset of items included in the questionnaire. The original scale is very extensive (it has more than 50 items) and it was not possible to include all items in our questionnaire.

Table 2.5: Descriptive statistics of outcomes at follow-up

	N (1)	Mean (2)	Std. Dev. (3)	Min (4)	Max (5)
Parental investment (quantity)					
Time					
Parental Time Investment Index	272	12.971	1.932	6.9	17.7
Involvement in physical games	338	4.191	0.785	1.9	6
Involvement in physical games every day	338	0.353	0.261	0	1
Involvement in didactic activities	358	4.825	0.989	1.8	6
Involvement in didactic activities every day	358	0.452	0.351	0	1
Involvement in socialization activities	318	3.879	0.875	1.8	6
Involvement in socialization activities every day	318	0.252	0.251	0	1
Father's involvement in childrearing	388	0.769	0.352	0	1
Material resources					
Toys	374	0.851	0.140	0.4	1
5 or more children books	389	0.843	0.364	0	1
Parental investment (quality)					
Positive Parenting Scale (E2P)	346	4.053	0.437	2.1	5
E2P: Attachment	374	4.340	0.585	1.1	5
E2P: Routines	378	3.820	0.917	1	5
E2P: Social support	378	3.755	0.998	1	5
E2P: Parental reflection	369	3.570	0.741	1	5
Violent disciplinary approach	382	0.390	0.488	0	1
Parental Stress					
Parental Stress Index (PSI)	260	73.019	17.755	37	147
PSI: Parental discomfort	348	26.949	8.362	12	56
PSI: Dysfunctional interaction	333	20.619	6.305	12	50
PSI: Difficult child	308	25.971	7.030	12	52
Parental Sense of Competence					
Parental Sense of Competence Scale (PSOC)	332	4.119	0.528	2.5	5.7
PSOC: Effectiveness	354	4.129	0.806	1.2	6
PSOC: Controllability	368	3.862	1.094	1	6
Knowledge					
Index of True/False questions	386	10.808	1.914	1	13
Time preferences					
Discount rate	332	0.081	0.089	0.0	0.25

Note: N= number of observations. Std. Dev.= Standard Deviation. Min= Minimum. Max= Maximum.

2.4.4 Estimation Approach

Our simplest specification uses OLS to regress Y_{ic}, the outcome of interest corresponding to family i in CAIF center c, on ITT_{ic}, an indicator taking the value of one if the family was randomly assigned to receive messages in a CAIF center selected to participate in the messaging intervention, and zero otherwise; and on $Spill_{ic}$, an indicator equal to one if the family was assigned not to receive messages in a CAIF center selected to participate in the messaging intervention, and zero otherwise:

$$Y_{ic} = \gamma_1 ITT_{ic} + \gamma_2 Spill_{ic} + \mu_{ic} \tag{2.1}$$

This specification accounts for the two-stage randomization structure, which was conducted to assess the degree of spillovers of the messaging intervention on CAIF families that were not receiving messages, but that attended an early childhood center in which other families were receiving messages.

We adjusted Equation 2.1 subsequently for the following set of covariates: (i) $Strata_{ic}$, the variable used to stratify the randomization at the CAIF center level (a dummy equal to one if the average level of education of mothers participating in *Crianza Positiva* at the center was at least middle school); (ii) B_{ic}, a set of covariates that were unbalanced after attrition at a significant level below 10% (the mother's age, whether the child lived with other adults aside from mother and/or father, the number of other children in the household); and (iii) X_{ic}, other covariates that could help improve the precision of the estimation (the child's age and gender, mother's age and education, whether the child lived with both biological parents, time elapsed since the messaging intervention begun, whether the family had a negative shock in the 12 months prior to the intervention, and a baseline measure of the outcome, when available). Our final specification was as follows:

$$Y_{ic} = \beta_1 ITT_{ic} + \beta_2 Spill_{ic} + \beta_3 Strata_c + B'_{ic}\beta_4 + X'_{ic}\beta_5 + \epsilon_{ic} \tag{2.2}$$

Assignment to the messaging treatment correlates strongly with receiving the messages (the correlation is 89% - see the first stage regression in Appendix Table 2.ii). Hence, because the local average treatment effect (LATE) can be calculated by dividing the ITT effect by the first stage, we should expect that the LATE would be similar to the ITT.

Because our experimental design involved a clustered randomization, we adjusted standard errors for the clustered design (Abadie et al., 2017). The usual approach when the number of clusters is large is to estimate standard errors using the Zeger and Liang (1986) covariance estimator. Unfortunately, the number of clusters in our data is only 24. As recommended by Cameron et al. (2008) we use a wild bootstrap with null imposed to enable more accurate cluster-robust inference. In addition, we conducted randomization inference which provides the means to assess whether the treatment effect estimate is unlikely to be observed by chance and is hence statistically significant (Heß, 2017). For this purpose, we observe the same coefficient estimate across 2000 different realizations of the treatment assignment. We report the associated p-values.

For each family of outcomes, we adjust standard errors to account for multiple hypotheses using the Romano-Wolf correction described in Romano and Wolf (2005a,b, 2016).[25] We exclude summary indices from these adjustments.

We also report the power of the sample to detect an effect of 0.25 standard deviations in the case of continuous outcomes and an effect of a 10 percentage points in the case of discrete outcomes. The power calculations account for the experimental design (randomization at two levels) and for intra-cluster (intra-CAIF center) correlation of the outcome.

2.5 Impact Evaluation: Results

2.5.1 Core Results

We begin by analyzing the results on families that were randomized to receive messages and focus on spillovers afterwards. Tables 2.6, 2.7 and 2.8 report the results of regressions of each outcome on an Intention to Treat (ITT) indicator and an indicator for whether the subject belongs to the spillover sample (Spill). In Table 2.6, we report results on parental quantity of investment with respect to time and material resources and in Table 2.7 we report results on the quality of this investment. In Table 2.8, we report results on parental stress, sense of parental competence, parental knowledge and discount factor.

The first column in each of these tables shows the raw treatment-control differences and the coefficient's unadjusted standard error. The second column shows ITT effects after adjusting for the stratum used for randomizing CAIF centers (average maternal education above middle school), child's gender and age in months, and maternal education. Column (3) adds the following covariates to the former regression: mother's age, time elapsed since the messaging intervention begun, whether the family had a negative shock in the 12 months prior to the intervention, number of other children in the household, whether the child lived with other adults in addition to father and mother, and whether the child lived with her biological father and mother (intact family). Column (4) adds a control for the outcome at baseline, if available.[26] While results are quite robust across specifications, we cen-

[25] This procedure uses resampling methods, to control for the probability of rejecting at least one true null hypothesis in the family of hypotheses under test. In addition to controlling for the familywise error rate, the methodology offers considerably more power and results in p-values that are always weakly smaller than those obtained from classical multiple testing procedures, such as Bonferroni and Holm (Clarke et al., 2020).

[26] To avoid losing observations due to missing values in the covariates, we generated, for each covariate, a dummy equal to one if the observation was missing and imputed the missing value

ter our discussion on that in Column (3), which improves precision and adjusts for the few imbalances in covariates across treatment arms. In Columns (5), (6) and (7) we report adjusted p-values corresponding to the outcome difference by ITT status reported in Column (3). Column (5) shows the p-values when adjusting for clustering, Column (6) shows the p-values when adjusting for multiple hypothesis testing (MHT) and Column (7) shows the p-values that result from randomization inference.

We also report in Column (8) the standard deviation of each outcome for the control sample (so that effects can be measured in terms of standard deviations) and in Column (9) the sample power to detect pre-established effect sizes. In particular, the power calculations consider an effect size of 0.25 standard deviations in the case of continuous outcomes and of 10 percentage points in the case of dichotomous outcomes. The calculations account for the experimental design (randomization at two levels) and for intra-cluster (intra-CAIF center) correlation of the outcome.

Parental Quantity of Investment: Time and Material Resources

The first outcome in Table 2.6 is the aggregate index of parental time investment, which shows an ITT effect of 0.33 standard deviations (obtained by dividing 0.667 by 2.020 – see Columns (3) and (8)). For comparison, Attanasio et al., 2018 evaluate a scalable parenting intervention in Colombia and find that an increase of 0.34 standard deviations in the home environment quality could work as suggestive mechanism for an increase of 0.15 standard deviations on overall child development.[27] Our effect is statistically significant when adjusted by the clustered design, by multiple hypotheses testing (MHT), and by randomization inference. We also find that the intervention increases the average frequency of parental engagement in didactic activities with the child by 0.29 standard deviations (0.299/1.024). The effect is significant at a 1% level when considering unadjusted p-values, and at the 5% level when adjusting for clustering (Column (5)), when adjusting for MHT (Column (6)), and when considering randomization inference (Column (7)). We find that the messaging intervention increases the frequency of parental involvement in physical games with the child in a magnitude of 0.27 standard deviations (0.217/0.794). This estimate is significant at a level of 5% when considering unadjusted p-values, when considering p-values that are adjusted for clustering, and when adjusting for MHT, and at the 10% significance level when conducting randomization inference. The ITT effect on the frequency of parental involvement in social activities with the

with the average of the covariate in the sample.

[27]The authors measure child development using the Bayley Scales of Infant and Toddler Development and home environment quality by combining information from the number of magazines, books or newspapers in the home, the number of toy sources, the number of varieties of play materials in the home and the number of play activities the child engaged in with adults.

child has a magnitude of 0.27 standard deviations (0.229/0.851) and is significant at the 5% level, but loses significance in the absence of controls (Column (1)). However, we find statistically significant effects in the frequency of involvement in social activities, didactic activities and physical games when we dichotomize the outcome and place a value of one to daily involvement in that activity and zero otherwise (See Appendix Table 2.v). We find no statistically significant ITT effects on father's involvement in childrearing nor on availability of toys or books in the household. The power to detect an effect of 0.25 standard deviations is 0.73 for the parental investment index, 0.53 for the measure of engagement in physical games, and 0.61 and 0.67 for engagement in didactic activities and social activities respectively.

Parental Investment Quality

Table 2.7 reports results on the quality of parental investment. We find a positive effect of 0.31 standard deviations (0.122/0.395) on the index of positive parenting when considering unadjusted p-values. The effect maintains statistical significance at the 1% level when we run randomization inference or when we account for the clustered sample design, and at the 5% level once we account for multiple hypotheses testing. Similarly, we find that the intervention has an effect on parental outreach for social support of 0.25 standard deviations (0.239/0.954) that is significant at the 5% level when considering unadjusted p-values, clustering (Column (5)) and randomization inference (Column (7)), and at the 10% after adjusting for MHT (Column (6)). Moreover, we find a statistically significant ITT effect (at the 1% level when considering either unadjusted p-values, clustering, or randomization inference, and at the 5% level after adjusting for MHT) on parents' reflective capacity. The magnitude of the effect is of 0.33 standard deviations (0.243/0.727). Power ranges around 0.5 for all outcomes.

Parental Stress, Sense of Competence, Discount Rate and Knowledge

In Table 2.8 we report the effects of the messaging intervention on parental stress, sense of competence, discount rate, and knowledge about positive parenting. We find no evidence of an effect of the intervention on parental stress, discount rate and parental knowledge. On the other hand, we find a positive and statistically significant effect on the sense of parental effectiveness when considering unadjusted p-values, but the effect becomes non-significant once we adjust standard errors for either MHT, clustering in the sample, or randomization inference. The power to detect a 0.25 standard deviation effect size is 52%. The power is lower in the case of the controllability outcome. In the latter outcome, we do not find a statistically significant effect.

Table 2.6: Intention to Treat (ITT) effects on parental quantity of investment

	(ITT=1)-(ITT=0)				p-value adjusting for clustering	p-value adjusting for MHT	p-value from randomization inference	ITT=0 Std. Dev.	Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Parental time investment									
Time Investment Index	0.488**	0.510**	0.667**	0.605**	0.006	0.010	0.011	2.020	0.730
	(0.246)	(0.247)	(0.247)	(0.233)					
Physical games	0.172*	0.175*	0.217**	0.234***	0.020	0.050	0.053	0.794	0.526
	(0.093)	(0.093)	(0.093)	(0.088)					
Didactic activities	0.230**	0.274**	0.299***	0.260***	0.020	0.020	0.012	1.024	0.606
	(0.107)	(0.108)	(0.109)	(0.097)					
Social activities	0.169	0.183*	0.229**	0.164*	0.026	0.050	0.037	0.851	0.666
	(0.103)	(0.104)	(0.103)	(0.098)					
Father's involvement	-0.022	-0.015	-0.029	-0.033	0.449	0.287	0.436	0.341	0.466
	(0.038)	(0.038)	(0.026)	(0.023)					
Material resources									
Toys	-0.026*	-0.024*	-0.021	n/a	0.168	0.248	0.193	0.128	0.408
	(0.015)	(0.014)	(0.014)						
5 or more children books	0.001	0.003	0.001	n/a	0.999	0.970	0.969	0.360	0.541
	(0.039)	(0.039)	(0.039)						
No controls	X								
Basic controls		X	X	X					
Basic and additional controls			X	X					
Basic, additional controls, and outcome at baseline if available				X					

Note: Rows depict different outcomes. Families of outcomes are identified by a heading in italics. Columns (1)-(4) report the ITT coefficient for different specifications. The estimations control for an indicator of whether the subject was randomized to control within treated CAIF centers. Our preferred specification is reported in Column (3). We adjust the p-value of the coefficient of our preferred specification for clustering (Column (5)), multiple hypotheses testing (Column (6)) and randomization inference (Column (7)). Column (8) reports the standard deviation of the outcome in the control sample and Column (9) indicates the sample power to detect an effect size of 0.25 standard deviations in the case of continuous outcomes and of 10 percentage points in the case of dichotomous outcomes. The calculations account for the experimental design (randomization at two levels) and for intra-cluster (intra-CAIF center) correlation of the outcome. Basic controls= Strata, child's age and gender, and mother's education. Additional controls= Mother's age, time elapsed since 1st message, negative shocks, other children in household, other adults in household and intact family. Std. Dev.= Standard Deviation. n/a=not available. Robust standard errors in parentheses. * p<0.1, ** p<.05, *** p<.01

Table 2.7: Intention to Treat (ITT) effects on parental quality of investment

	(ITT=1)-(ITT=0)				p-value adjusting for clustering	p-value adjusting for MHT	p-value from randomization inference	ITT=0 Std. Dev.	Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Parental quality of investment									
Positive Parenting Scale (E2P)†	0.098*	0.106**	0.122**	n/a	0.003	0.040	0.003	0.395	0.531
	(0.050)	(0.050)	(0.049)						
E2P: Attachment	0.062	0.102*	0.058	n/a	0.110	0.475	0.185	0.558	0.458
	(0.065)	(0.052)	(0.051)						
E2P: Routines	0.116	0.131	0.093	n/a	0.340	0.475	0.332	0.925	0.626
	(0.098)	(0.096)	(0.093)						
E2P: Social support	0.222**	0.225**	0.239**	n/a	0.016	0.089	0.018	0.954	0.457
	(0.106)	(0.104)	(0.103)						
E2P: Parental reflection	0.195**	0.215**	0.243***	n/a	0.003	0.020	0.006	0.727	0.553
	(0.082)	(0.083)	(0.084)						
Violent disciplinary approach	-0.011	-0.012	-0.032	n/a	0.622	0.535	0.677	0.486	0.100
	(0.053)	(0.054)	(0.054)						
No controls	X								
Basic controls		X	X	X					
Basic and additional controls			X	X					
Basic, additional controls, and outcome at baseline if available				X					

Note: See Note in Table 2.6. †The E2P index and subscales do not include the full set of E2P items in the original scale.

Table 2.8: ITT effects on parental stress, sense of competence, discount rate & knowledge

	(ITT=1)-(ITT=0)				p-value adjusting for clustering	p-value adjusting for MHT	p-value from randomization inference	ITT=0 Std. Dev.	Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Parental Stress									
Parental Stress Index (PSI)	1.572	0.919	0.973	2.702	0.818	0.604	0.702	16.277	0.276
	(2.249)	(2.158)	(2.188)	(2.028)					
PSI: Parental discomfort	1.432	1.360	1.022	1.610*	0.471	0.515	0.468	7.979	0.305
	(0.920)	(0.903)	(0.924)	(0.836)					
PSI: Dysfunctional interaction	-0.355	-0.444	-0.447	-0.380	0.544	0.515	0.607	5.982	0.373
	(0.739)	(0.735)	(0.730)	(0.712)					
PSI: Difficult child	0.798	0.932	1.059	1.411**	0.389	0.455	0.289	5.905	0.385
	(0.775)	(0.784)	(0.783)	(0.709)					
Parental Sense of Competence									
Parental Sense of Competence Scale (PSOC)	-0.023	-0.011	-0.002	-0.017	0.968	0.604	0.980	0.519	0.538
	(0.062)	(0.063)	(0.064)	(0.058)					
PSOC: Effectiveness	0.173*	0.176**	0.157*	0.135	0.132	0.168	0.189	0.795	0.516
	(0.089)	(0.087)	(0.095)	(0.087)					
PSOC: Controllability	-0.169	-0.149	-0.145	-0.192*	0.547	0.248	0.442	1.048	0.380
	(0.119)	(0.119)	(0.119)	(0.114)					
Information									
Parental knowledge	-0.283	-0.228	-0.271	n/a	0.334	0.168	0.363	1.715	0.347
	(0.201)	(0.187)	(0.193)						
Time preferences									
Discount rate	0.009	0.010	0.009	0.000	0.416	0.376	0.454	0.088	0.372
	(0.010)	(0.010)	(0.010)	(0.010)					
No controls	X								
Basic controls		X	X	X					
Basic and additional controls			X	X					
Basic, additional controls, and outcome at baseline if available				X					

Note: See Note in Table 2.6.

2.5.2 Heterogeneity and Mechanisms

To assess whether the messaging program operates through behavioral channels, we explore program heterogeneity across three dimensions of parental preferences, constraints, and beliefs measured prior to the intervention: the parental discount rate, negative shocks faced by the household in the previous 12 months, and parental sense of competence. Our baseline assessment suggested that behavioral barriers would be stronger in parents with high discount rates, families with more negative shocks in the past, and parents with low sense of competence. Because the intervention was designed to address these barriers, we expected it to have stronger effects among these families.

Results are presented in Tables 2.9 and 2.10. Each triplet of columns correspond to a different behavioral barrier. The first two columns within each triplet show the coefficients and standard errors from an OLS regression of the outcome in each row on the ITT main effect, the interaction between ITT status and the variable capturing the behavioral barrier, the main effect of the behavioral barrier, maternal education, government assistance, and randomization strata. The first column displays the ITT main effect and the second column the interaction of ITT with the behavioral barrier. In the third column of each triplet we report the total effect (main effect + interaction effect) and its associated standard error. Column (2) of Table 2.9 shows that families with higher cognitive fatigue (higher likelihood of at least two negative shocks in the past 12 months) derive some additional benefits from the intervention.[28] In particular, families facing a larger number of negative shocks at baseline are more likely to implement routines and are less likely to use violent discipline. Moreover, the total effect in the latter cases is significant (see Column (3)).

Column (5) explores whether parents with lower initial parental sense of competence benefit more from an intervention geared towards providing encouragement and constructing positive identities and Column (6) shows whether the total effect is significant for this group. Parents with low initial self-esteem improve their scores in several measures of quality of investment, including establishing routines and reflecting on parenthood. The total effect is significant for the outcomes relating to the quality of parental investment, but not for the number of toys.

Results in Column (8), on the other hand, show no evidence that the intervention was more effective among parents experiencing higher discount rates.

Table 2.10 shows that families with higher cognitive fatigue have a stronger

[28]Treatment and control families do not differ in the mean values for these variables. We also confirmed that other baseline socioeconomic variables were balanced within the samples defined by the dichotomous variables used in the heterogeneity analysis (balance analysis is available upon request).

sense of parental competence at follow-up (see Column (2)). In the parental sense of competence scale and the subscale of effectiveness, the total effect for those with a higher number of negative shocks in the past 12 months is positive and significant (see Column (3)).

In sum, we find evidence that the program has stronger effects on families with initially higher exposure to cognitive fatigue and more negative identities. These differential effects operate mainly on the qualitative margin. By suggesting simple activities to carry out at home, parents are more able to establish routines, need to rely less on the use of violent discipline and their sense of parental competence increases. The encouragement provided to parents through the messages seems to improve their reflection capacity, their ability to organize their routines, and material investments.

Table 2.9: Heterogeneity on parental investment by behavioral barriers

	Cognitive fatigue			Negative identity			Present bias		
	ITT	ITT × Negative shocks	Total effect	ITT	ITT × Low parental efficacy	Total effect	ITT	ITT × High discount rate	Total effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Investment Quantity									
Time Investment Index	0.557*	-0.117	0.440	0.563	0.463	1.026**	0.614*	-0.332	0.282
	(0.307)	(0.513)	(0.418)	(0.476)	(0.671)	(0.514)	(0.324)	(0.554)	(0.470)
Physical games	0.257**	-0.271	-0.014	0.042	0.334	0.376**	0.148	0.098	0.246
	(0.115)	(0.184)	(0.150)	(0.196)	(0.256)	(0.181)	(0.118)	(0.195)	(0.164)
Didactic activities	0.266**	-0.005	0.262	0.019	0.305	0.324	0.266*	0.062	0.329*
	(0.135)	(0.221)	(0.179)	(0.203)	(0.288)	(0.217)	(0.146)	(0.237)	(0.192)
Social Activities	0.183	0.048	0.231	0.406*	-0.150	0.256	0.309**	-0.233	0.076
	(0.131)	(0.210)	(0.167)	(0.219)	(0.312)	(0.233)	(0.134)	(0.227)	(0.188)
Father's involvement	0.002	-0.036	-0.033	0.018	-0.079	-0.061	0.005	-0.055	-0.050
	(0.043)	(0.080)	(0.069)	(0.079)	(0.102)	(0.068)	(0.050)	(0.079)	(0.063)
Toys	-0.018	0.005	-0.013	-0.057*	0.083**	0.026	-0.024	0.031	0.007
	(0.017)	(0.030)	(0.025)	(0.030)	(0.040)	(0.028)	(0.017)	(0.033)	(0.028)
5 or more children books	-0.010	0.029	0.019	0.065	-0.099	-0.034	-0.049	0.091	0.043
	(0.046)	(0.079)	(0.067)	(0.087)	(0.113)	(0.079)	(0.050)	(0.087)	(0.073)
Investment Quality									
Positive Parenting Scale (E2P)	0.089	0.064	0.154*	-0.003	0.290*	0.287***	0.101*	-0.028	0.074
	(0.060)	(0.101)	(0.084)	(0.113)	(0.147)	(0.098)	(0.061)	(0.114)	(0.098)
E2P: Attachment	0.067	0.044	0.111	0.089	0.210	0.299*	0.061	0.014	0.075
	(0.055)	(0.098)	(0.085)	(0.093)	(0.183)	(0.171)	(0.058)	(0.110)	(0.097)
E2P: Routines	-0.021	0.411**	0.389**	-0.085	0.487*	0.401*	0.119	-0.103	0.016
	(0.117)	(0.191)	(0.156)	(0.188)	(0.272)	(0.211)	(0.120)	(0.206)	(0.176)
E2P: Social support	0.188	0.124	0.312*	-0.103	0.476	0.373*	0.224*	0.045	0.270
	(0.132)	(0.206)	(0.164)	(0.227)	(0.313)	(0.225)	(0.130)	(0.231)	(0.197)
E2P: Parental reflection	0.187*	0.051	0.238*	-0.033	0.458*	0.425**	0.195*	0.037	0.233
	(0.103)	(0.162)	(0.129)	(0.186)	(0.247)	(0.173)	(0.104)	(0.176)	(0.146)
Violent disciplinary approach	0.053	-0.234**	-0.180**	0.084	-0.175	-0.091	-0.072	0.184	0.113
	(0.065)	(0.104)	(0.085)	(0.108)	(0.148)	(0.103)	(0.070)	(0.114)	(0.093)

Note: Table reports OLS regressions. Each triplet of columns correspond to a different behavioral variable. Cognitive fatigue is measured with an indicator that takes value of 1 if the family was exposed to 2 or more negative shocks in the previous 12 months to the survey. Negative identity is measured with an indicator that takes the value 1 if the parental efficacy at baseline is less or equal than 4 (ranges from 1 to 6). Present bias is proxied with the discount rate. Each row denotes a different outcome. For each outcome, we show the coefficient and standard error for the ITT main effect (first column in each triplet), the coefficient and standard error for the interaction between ITT and the behavioral barrier analyzed (second column in each triplet), and the total effect in each group (third column in each triplet). The estimations control for an indicator of whether the subject was randomized to control within treated CAIF centers (the "spillover" sample). All regressions adjust, in addition, for randomization strata, maternal education, and governmental assistance. Standard errors in parentheses. * p<0.1, ** p<.05, *** p<.01

Table 2.10: Heterogeneity on parental stress, sense of competence and knowledge by behavioral barriers

	Cognitive fatigue			Negative identity			Present bias		
	ITT	ITT × Negative shocks	Total effect	ITT	ITT × Low parental efficacy	Total effect	ITT	ITT × High discount rate	Total effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Parental Stress Index (PSI)	0.464	1.006	1.470	2.363	-0.202	2.161	1.693	-0.096	1.598
	(2.610)	(4.351)	(3.543)	(4.487)	(5.911)	(4.187)	(2.874)	(4.389)	(3.391)
PSI: Parental discomfort	1.822*	-1.111	0.711	2.280	-0.063	2.218	0.980	1.656	2.637*
	(1.068)	(1.795)	(1.480)	(1.761)	(2.486)	(1.825)	(1.184)	(1.868)	(1.473)
PSI: Dysfunctional interaction	-0.822	0.621	-0.201	-1.193	2.376	1.183	-0.765	0.619	-0.146
	(0.806)	(1.562)	(1.364)	(1.350)	(1.955)	(1.524)	(0.953)	(1.410)	(1.092)
PSI: Difficult child	0.601	-0.082	0.519	1.366	-0.138	1.228	1.658	-0.952	0.705
	(0.929)	(1.561)	(1.277)	(1.677)	(2.218)	(1.526)	(1.040)	(1.594)	(1.237)
Parental Sense of Competence Scale (PSOC)	-0.121	0.294**	0.173*	0.051	-0.099	-0.048	-0.086	0.196	0.110
	(0.079)	(0.119)	(0.093)	(0.142)	(0.187)	(0.132)	(0.080)	(0.133)	(0.109)
PSOC: Effectiveness	0.015	0.347**	0.362***	0.226	-0.034	0.192	0.070	0.222	0.292*
	(0.111)	(0.169)	(0.132)	(0.179)	(0.264)	(0.195)	(0.114)	(0.186)	(0.153)
PSOC: Collaboration	-0.349**	0.489**	0.140	-0.059	-0.111	-0.170	-0.185	0.014	-0.171
	(0.141)	(0.236)	(0.193)	(0.257)	(0.337)	(0.239)	(0.155)	(0.257)	(0.207)
Parental knowledge	-0.073	-0.258	-0.331	0.067	-0.169	-0.102	-0.116	0.210	0.094
	(0.227)	(0.399)	(0.335)	(0.361)	(0.502)	(0.377)	(0.213)	(0.369)	(0.312)

Note: See Note in Table 2.9.

2.5.3 Spillovers

After participating in the workshop, parents could potentially interact in meetings, local sites, or social media. To assess spillovers, we compared the effects of non-treated families within treated centers with families within non-treated centers.

Table 2.11 shows the coefficients on an indicator equal to one if the family was assigned not to receive messages in a CAIF center assigned to treatment, and zero if the family belonged to a CAIF center assigned to control (β_2 in Equation 2.2 when including the following controls: strata, the child's age and gender, mother's age and education, whether the child lived with both biological parents, whether the child lived with other adults aside from mother and/or father, the number of other children in the household, time elapsed since the messaging intervention begun, and whether the family had a negative shock in the 12 months prior to the intervention). Any positive effect of the intervention on this indicator would suggest spillover effects. Each row corresponds to a different outcome. Overall, the evidence is not supportive of the idea of positive spillovers. There are only two statistically significant effects which are robust to familywise multiple hypothesis testing adjustment, but the statistical power is quite low.[29] Moreover, the effects also run in the opposite way than hypothesized. Families in this group show higher levels of violent discipline and lower levels of knowledge of positive parenting competences. The results could express frustration among the untreated parents in treated centers for excluding them from participating.

[29]Power estimations consider the outcome's standard deviation in each sample, which may explain why, for the case of some outcomes, we find a higher power in the spillover analysis than in the main analysis.

Table 2.11: Spillover effects

	Coeff. and s.e.	MHT adjusted p-value	Power
Time Investment Index	0.188		0.587
	(0.386)		
Physical games	0.188		0.574
	(0.123)		
Didactic Activities	0.019		0.502
	(0.187)		
Social Activities	0.146		0.602
	(0.168)		
Father's involvement	-0.025		0.308
	(0.043)		
Toys	-0.019		0.429
	(0.026)		
5 or more children books	-0.035		0.528
	(0.067)		
Positive Parenting Scale (E2P)	0.011		0.583
	(0.075)		
E2P: Attachment	0.033		0.629
	(0.079)		
E2P: Routines	0.236		0.559
	(0.171)		
E2P: Social support	-0.293*	0.238	0.461
	(0.171)		
E2P: Parental reflection	0.080		0.571
	(0.120)		
Violent disciplinary approach	0.170**	0.119	0.095
	(0.082)		
Parental Stress Index (PSI)	5.065		0.349
	(3.676)		
PSI: Parental discomfort	3.030*	0.149	0.345
	(1.541)		
PSI: Dysfunctional interaction	-0.038		0.380
	(1.030)		
PSI: Difficult child	1.471		0.447
	(1.258)		
Parental Sense of Competence Scale (PSOC)	-0.057		0.502
	(0.094)		
PSOC: Effectiveness	0.051		0.449
	(0.152)		
PSOC: Controllability	-0.126		0.411
	(0.184)		
Parenting knowledge	-0.774**	0.050	0.486
	(0.323)		
Time discount rate	0.015		0.419
	(0.017)		

Note: Table reports β_2 in Equation 2.2 when including the following controls: strata, the child's age and gender, mother's age and education, whether the child lived with both biological parents, whether the child lived with other adults aside from mother and/or father, the number of other children in the household, time elapsed since the messaging intervention begun, and whether the family had a negative shock in the 12 months prior to the intervention. N corresponds to number of observations. Standard errors in parentheses. Table reports coefficient on a dummy indicating that the family was assigned to the control group in a center assigned to treatment. Coeff.= Coefficient. s.e.= Standard errors. Standard errors in parentheses. * $p<0.1$, ** $p<.05$, *** $p<.01$

2.6 Conclusions

This chapter evaluates the impact of a text and audio messaging program (a component of *Crianza Positiva*) aimed at helping parents develop and sustain parenting competences over time. The program reminds parents about the benefits of engaging in positive parenting practices, provides them with suggestions of simple and concrete positive parenting activities, reinforces positive parental identities, and encourages parents to seek resources within their families and community to improve their parenting behaviors and attitudes.

This chapter, is among the first, together with York et al. (2019), to show the benefits of a program combining e-messaging and nudges in boosting parental investment; and the first we are aware of that implements these tools to address parenting in a developing country. Unlike other parenting programs using e-messages, our program covers a comprehensive range of parenting areas, including sensitive observation and response, the importance of a safe and nurturing environment, the importance of speaking and reading to the child, the key role of free play, and the value of self-caring and of having a reflective parenting attitude. Moreover, our program is focused on parents with children aged 0-2, while most of the literature focuses on parents of older children. Also, our evaluation includes an extensive set of outcomes which are measured using validated instruments. Unlike previous literature, we not only assess the quantity but also the quality of parental investment, including parental stress, parental sense of competence, parental knowledge about parenting, and parental sensitivity. The program complemented a prior 8-week workshop for families at local early childhood centers, and sent caregivers text and voice messages through their cellphones (via SMS and WhatsApp, respectively). Families received both types of messages three times a week during 24 weeks between January and June 2018.

The program was well-received by families. Among families assigned to treatment, 95% said that the messages had been either very useful (61%) or somehow useful (34%), and only one family opted out of the audio messages. Furthermore, the program had a positive effect on the frequency of parental involvement with the child, on parental competences and parenting attitudes. Our findings show that messages had an impact of 0.33 standard deviations on a parental time investment index and on parental engagement in social, physical, and didactic activities with the child. They also increased parents' quality of investment as measured by a positive parenting index and by an index of outreach for social support (by 0.31 and 0.25 standard deviations respectively). Moreover, the messages improved parental capacity to reflect on parenting by 0.33 standard deviations.

We find that, for several outcomes, the program had stronger effects over parents with initial negative identity (i.e., a low sense of parental competence) and over parents experiencing negative shocks in the months prior to the intervention (higher cognitive fatigue). Because the intervention targeted these behavioral barriers, our results suggest that it may have triggered the right channels. We could also infer from these findings that nudges involving recommendations of simple actions and messages of encouragement may be particularly important to activate changes in parental decision-making.

Overall, our results indicate that the combination of e-messages and nudges are a promising tool to enhance parental behaviors, competences and attitudes. Our findings indicate that the program is highly cost-effective: the cost per family of implementing this program is US$ 5 if only SMS were to be used, and even less if only WhatsApp messages were to be considered. In terms of external validity, our intervention was implemented in CAIF centers in Uruguay, which tend to assist families of lower socioeconomic status, but the program was designed for any socioeconomic setting and would need little adaptation to be delivered in other contexts. Moreover, the intervention was embedded directly in a governmental-provided program, which made the implementation and results close to a "real-life" intervention.

While the self-reported nature of our outcomes could be pointed out as a caveat, we don't believe our findings are being led by social desirability bias. All parents in the treatment and control groups participated in an 8-week parenting workshop prior to the implementation of the messaging intervention, focused around the key competences associated with positive parenting. Indeed, while in Balsa et al., 2020 we find that the workshop itself had positive effects on an index of parenting knowledge about positive parenting competences, we find no impact of the messaging intervention on the same outcome. Furthermore, our endline measurements include mostly internationally validated scales which have been used in multiple and diverse settings, reducing concerns about the messaging program being designed to "teach to the test".

The fact that the *Crianza Positiva* messaging program was implemented right after families had finished the parenting workshop raises another issue. We cannot affirm that the messages would have been effective in the absence of this workshop. However, the literature shows that short group-based parenting interventions have a hard time consolidating and sustaining effects over time, and home visits are hard to maintain due to their high costs. At a minimum, our study underscores the value of messages in helping integrate and put into practice concepts acquired in prior face-to-face parenting interventions. This is supported by findings in Balsa et al., 2020, that show that the workshop by itself does not increase the frequency of parental involvement with the child. In addition, many RCTs are conducted on top of interventions that already exist, and usually results are hard to disentangle completely from the root program.[30] As such, our study offers suggestive evidence about the potential that e-messaging programs could have as stand-alone. In future research, we would like to obtain administrative data with objective child outcomes and to merge this information with our database. Up to now, we were not able to collect child development outcomes due to budget constraints.[31] We also plan to explore the effectiveness of messages without the requirement of participating in a previous workshop, to assess whether families' responses to the messages vary according to the sender (sender versus content effect), and to improve our understanding of the mechanisms behind the effects by randomly assigning message types.

[30]For example, York et al. (2019) study the effectiveness of a texting message to parents of preschoolers. Their program is placed on top of an already existing preschool program and shares the foundations of the school curriculum.

[31]We have evaluated the effect of the *Crianza Positiva* e-messaging program on caregiver-child language interaction patterns (see Balsa et al. (2021)). Our outcomes are externally assessed rather than self-reported, reducing the potential incidence of desirability bias. We find that the intervention was successful at improving the quality of parental vocalizations, as measured by the parent's pitch range. We also found suggestive evidence of increases in the duration of adult vocalizations. These findings are consistent with more frequent parental involvement, as we find in this chapter.

2.A Policies to Counteract Socioeconomic Gaps in Early Childhood: Parenting Programs

Yeung et al. (2002) highlight two reasons why families of low socioeconomic background show, on average, lower parental investment levels: (i) lower income available to purchase materials, experiences and services that contribute to the development of children's human capital, and (ii) different family processes. Vulnerable families have fewer resources to invest in education, health, food, housing, child stimulation material, and toys, among others. In addition, poverty can have an impact on the emotional state of the adults in the family and, hence, on their ability to interact with their children. Heckman (2006) argues that the lack of early stimulation is more important than the lack of economic resources in explaining developmental gaps in early childhood. Cunha (2015) develops a model in which a child's human capital is determined by the interaction between investments (for example, number of books at home) and institutions (for example, quality of school). Parents can adopt either a "concerted cultivation" parenting style (active engagement with institutions for the child's benefit) or a "natural growth" (passive role) style. Adopting the former has a cost, but the latter is costless. As parents from lower socioeconomic backgrounds are more income constrained, the model predicts they are more likely to adopt a "natural growth" parenting style.

A vast literature shows evidence of the effectiveness of parental interventions. In developed countries, these programs have shown positive effects on the development of children of around 0.3 to 0.5 standard deviations, as well as improvements in parental skills (Bakermans-Kranenburg et al., 2003; Blok et al., 2005; Al et al., 2014). Also in developing countries, parental interventions seem to be viable and effective mechanisms to improve the parent-child relationship, the parents' knowledge of the child's development and the child's mental and motor development (Nores and Barnett, 2010; Baker-Henningham and López Bóo, 2010; Knerr et al., 2013).

Home visits are among the most effective type of family intervention in low-income households. Home visits seek to improve parenting practices through the identification and assessment of existing resources in the household and their link with community resources. There is evidence that home visits improve parenting skills (Wilson, 2010), as well as children's cognitive abilities (Walker et al., 2015; Attanasio et al., 2014), socio-emotional abilities, and behavior (Pickering et al., 2014).[32] One limitation of home visiting programs is their high cost of scaling up.

[32] In Latin America, several of these programs have proven highly effective in reducing gaps in child development. A program in Jamaica, implemented between 1986 and 1989 (Grantham-McGregor et al., 1991; Gertler et al., 2014; Walker et al., 2011) was able to substantially improve children's cognitive development in the short run (0.8 standard deviations in 24 months) and influence the trajectory of cognitive skills, education, wages, and mental health up to 20 years

These programs require an intensive use of qualified facilitators, as well as high levels of training and supervision to ensure adequate execution of the intervention protocols (Leer et al., 2016). The intensity of these programs makes it difficult to reach broad sectors of the population that can also benefit from parental education. The literature has also found that the effects of some of these programs fade out over time (see overview by Bailey et al., 2017).

2.B Behavioral Barriers to Parental Investment and Tools to Overcome Them

Present bias

People have an order of preferences over choices in the long run, but when the time comes to make the decision, that order is reversed, usually giving up future benefits in order to obtain immediate rewards.[33] The evidence shows that most people tend to do less than optimal in a specific activity when the reward for that activity is received later (Thaler and Sunstein, 2009). There is also evidence that low socioeconomic status individuals discount the future at higher rates (Lawrance, 1991). Present-biased families are thus less likely to invest in costly activities that provide benefits only in the future (Agee and Crocker, 1996; Pabilonia and Song, 2013).

Commitments, reminders and immediate benefits are effective tools to overcome present bias. Commitments motivate people to be consistent with their objectives and increase the likelihood that this behavior will finally be carried out by imposing a psychological or material cost in case of non-compliance (Giné et al., 2010; Mayer et al., 2018). Reminders increase the salience of the future benefits of certain behaviors (Cunha et al., 2017). Text messages (York et al., 2019) are the most common and proven way to send reminders.[34] They have been used successfully in savings and weight loss programs (Karlan et al., 2016; Patrick et al., 2009).[35]

after its implementation. In Colombia, a similar program showed an increase in cognitive and receptive language development of 0.26 and 0.22 standard deviations respectively (Attanasio et al., 2014; Rubio-Codina et al., 2015). In Ecuador, Rosero and Oosterbeek (2011) find that home visits improved children's language skills by 0.4 standard deviations, memory performance by 0.6 standard deviations, and fine motor skills by 0.9 standard deviations. In Uruguay, Marroig et al. (2017) find that the home visiting program "Uruguay Crece Contigo" enhanced children's gross motor skills.

[33] This happens even in cases in which the short run benefits of waiting are quite large. In a famous experiment, Mischel et al. (1972) show that a group of children could not resist the temptation of eating a sweet, despite large benefits from waiting a few minutes.

[34] Other programs have used different channels to send messages. For example, Chong et al. (2016) show that media messages are effective to encourage adolescents to take supplemental iron pills.

[35] Other programs have sought to address present bias by providing immediate benefits from

Inattention and cognitive fatigue

Mullainathan and coauthors (Schilbach et al., 2016; Mani et al., 2013; Shah et al., 2012) argue that in poverty, the preoccupations that appear due to the scarcity of economic resources reduce the idle cognitive capacity and shorten the "bandwidth" available to make accurate decisions (Mani et al., 2013). This results in choices that are made quickly, intuitively and automatically, and that are more likely to fall into biases and errors. To overcome the sub-optimality of decisions due to cognitive fatigue, the literature proposes the use of reminder messages that make more salient the commitment to the desired objective (Bryan et al. (2010)) and the design of solutions that facilitate parenting practices by decomposing complex tasks into simpler ones (Mayer et al. (2018) and York et al. (2019)).

Negative identities

Negative identities may lead parents to opt for sub-optimal parental investments. When skills and effort are complements, an optimistic view of one's abilities can increase effort through higher levels of motivation (Bénabou and Tirole, 2002). In the context of parenthood, families need to trust in their capacity of influencing the trajectory of their children and must believe that their efforts are worthwhile. Identities are also related to the social group to which the individual belongs. The utility of parental investments can come not only from personal benefits, but also from how consistent that investment is in relation to what other members of the social group the individual belongs to also do (Akerlof and Kranton, 2000). Knowing what other parents in similar situations are doing can be useful as a reference point for deciding how to act. Increasing the salience of a positive identity can change both the way in which individuals evaluate their options and their performance (Gennetian et al., 2016).

Status-quo bias

Even knowing that changing parenting practices could be beneficial for a child's development, parents might find it costly to change their habits. Many of these habits reproduce the parenting patterns of prior generations. An effective way to overcome this bias is to set options by default (Madrian and Shea, 2001).

actions that will have a benefit in the future. Fryer Jr et al. (2015) show that providing monetary incentives for attendance and fulfillment of tasks in a parental program has positive effects on children's cognitive and non-cognitive test scores.

2.C Appendix Tables and Figures

The tables and figure included in this section supplement the information in the main text. The first table shows how different behavioral biases correlate with parental investment outcomes. The second table reports a first stage regression where the outcome variable is an indicator for whether the family actually received messages and the independent variable is the assignment to treatment (or intention to treat) status. Following we show a figure with the timeline of the project. The third table explains in detail how each evaluation measure is constructed. The fourth table compares descriptive statistics of the main sample with children aged 0-36 months in Uruguay. The fifth table reports intention to treat effects on a dichotomous indicator of every-day parental involvement.

Table 2.i: Regressions of parental investment outcomes on the discount rate, parental stress, and sense of competence

	Physical games	Stimulating activities	Social activities
Discount rate	0.133	-1.539**	0.017
	(0.568)	(0.729)	(0.670)
PSI: Dysfunctional interaction	-0.021**	-0.028**	-0.014
	(0.009)	(0.011)	(0.010)
PSOC	0.275**	0.232	0.257*
	(0.117)	(0.150)	(0.138)
Constant	3.311***	4.311***	3.127***
	(0.555)	(0.712)	(0.655)
N	289	289	289
r2	0.054	0.059	0.025
F	0.001	0.001	0.065

Note: Table shows results of regressing each parental investment outcome on the discount rate, parental stress and the sense of competence. PSI= Parenting Stress Index. PSOC= Parental Sense of Competence. Standard errors in parentheses. * p<0.1, ** p<.05, *** p<.01

Table 2.ii: Regression of an indicator of whether the family received any message on ITT status

	Family receives text messages
ITT status	0.886***
	(0.019)
Constant	-0.000
	(0.013)
N	529

Note: Standard errors in parentheses. * p<0.1, ** p<.05, *** p<.01

Figure 2.i: Timeline of the project

Table 2.iii: Construction of evaluation measures

Instrument	Construction of evaluation measure
Quantity of parental investment	
Parental engagement scale, by Cabrera et al. (2004)	The scale includes 32 items that were divided into four categories: physical games with the child (7 items), taking care of the child at home (8 items), didactic activities with the child (7 items), and socialization activities with the child (10 items). Respondents had to report their frequency of involvement in each task on a scale ranging from one to six. The lowest value corresponds to never getting involved and the highest to getting involved several or all days of the week. We averaged scores in each dimension and built a general parental investment score averaging all of them. For robustness, we also constructed dummy variables that took the value one if the respondent reported engaging in a certain activity every day and zero otherwise.
Father's Involvement subscale of the Etxadi-Gangoiti Scale	We included 9 out of the 11 items of the original scale. Two items were excluded because they were not applicable to the age group of the children under evaluation. For each item, the respondent had to answer whether the father of the child was regularly involved in certain activity. The results were coded as one or zero, with one being the equivalent of a "yes" answer. The total score was the summation of the answers to the nine questions.

Continued on next page

Table 2.iii – *Continued from previous page*

Instrument	Construction of evaluation measure
Quality of parental investment	
Positive Parenting Scale (E2P)	We selected seven items from the parental attachment subscale, four items from the reflective capacity subscale, and four items from the safety and stimulation subscales, from which we constructed an index of attachment and outreach for social support and another one indicating the degree to which parents could organize the child's activities around a daily routine. Respondents had to report their degree of agreement with several statements on a Likert scale ranging from one to five.
UNICEF MICS6 questionnaire	We constructed a dummy variable equal to one when the parent reported shaking, slapping, hitting the child, shouting at her, or calling her "silly" or "useless" during the past month.
Parental discount rates	
MCQ	The questionnaire includes 27 questions with binary options of an amount of money to receive today or a larger amount of money to receive at some point in the future. For example: "Which of these two options would you prefer: $ 1512 today or $ 1540 in 117 days?". Individuals were asked to choose one of two options for each question. The instrument identifies a time discount rate for each individual. There exists a value k that represents the point at which the respondent is indifferent between the two rewards. The values of the time discount rate range from 0 to 0.249, and a higher value indicates a higher preference for the present. In the follow-up survey, overall consistency is above 75% for 97% of respondents.

Continued on next page

Table 2.iii – *Continued from previous page*

Instrument	**Construction of evaluation measure**
Parental stress	
PSI/SF	The PSI/SF consists of 36 statements to which parents must respond on a Likert scale, with one being the lowest value and five the highest. The scale is divided into three subscales of 12 items each. The summation of the scores in each subscale determines the individual's total level of parental stress and varies between 36 and 180, with a higher score indicating higher levels of parental stress.
Parental sense of competence	
PSOC	The PSOC is a 16-item instrument in which the parent or caregiver classifies responses according to the degree of agreement with various statements. Each item is scored on a Likert scale that takes values between one and six, where one represents total disagreement with the proposed statement and six complete agreement. We constructed two subscales, (i) effectiveness, and (ii) controllability. Responses were averaged out for each subscale. A higher score is associated with a higher sense of parental competence.
Parental knowledge about positive parenting	
Index of parental knowledge	We constructed an index of parental knowledge as the count of the items responded correctly by the parent.

Table 2.iv: Comparison of descriptives statistics of main sample with Uruguayan National Survey of Nutrition, Child development and Health

	Main sample			Children aged 0-36 months that attend CAIF			Children aged 0-36 months		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's age	393	29.158	6.648	308	28.549	6.564	1647	30.151	6.803
Female child	397	0.491	0.501	313	0.479	0.500	1656	0.525	0.500
Child's age (in months)	393	23.962	6.525	313	21.284	9.130	1656	16.118	11.229
Intact family	391	0.752	0.432	308	0.698	0.460	1648	0.817	0.387
Beneficiary of cash transfers	397	0.662	0.473	313	0.719	0.450	1656	0.517	0.500
Only child	383	0.381	0.486	313	0.393	0.489	1656	0.386	0.487
Mother completed middle school	393	0.326	0.469	308	0.357	0.480	1643	0.445	0.497
Mother completed high school	393	0.295	0.457	308	0.260	0.439	1643	0.270	0.444

Note: Table shows descriptives statistics of the main sample (Columns (1)-(3)), children aged 0-36 months that attend CAIF centers in Uruguay (Columns (4)-(6)) and children aged 0-36 months in Uruguay (Columns (7)-(9)). Data corresponding to Columns (4)-(9) belongs to the Uruguayan National Survey of Nutrition, Child development and Health from 2018. N=Number of observations. Std. Dev.=Standard Deviation.

Table 2.v: Intention to Treat (ITT) effects on every-day parental involvement (dichotomous indicator)

	(ITT=1)-(ITT=0)				p-value adjusting for clustering	p-value adjusting for MHT	p-value from randomization inference	ITT=0 Std. Dev.	Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Physical games every day	0.071**	0.067**	0.074**	0.076**	0.056	0.089	0.061	0.251	0.396
	(0.030)	(0.031)	(0.031)	(0.030)					
Didactic activities every day	0.090**	0.093**	0.104**	0.076**	0.081	0.099	0.047	0.356	0.484
	(0.039)	(0.040)	(0.041)	(0.038)					
Social activities every day	0.067**	0.068**	0.073**	0.050*	0.081	0.099	0.039	0.245	0.632
	(0.030)	(0.030)	(0.032)	(0.029)					
No controls	X								
Basic controls		X	X	X					
Basic and additional controls			X	X					
Basic and additional controls, and outcome at baseline if available				X					

Note: Rows depict different outcomes. Families of outcomes are identified by a heading in italics. Columns (1)-(4) report the ITT coefficient for different specifications. The estimations control for an indicator of whether the subject was randomized to control within treated CAIF centers. Our preferred specification is reported in Column (3). We adjust the p-value of the coefficient of our preferred specification for clustering (Column (5)), multiple hypotheses testing (Column (6)) and randomization inference (Column (7)). Column (8) reports the standard deviation of the outcome in the control sample and Column (9) indicates the sample power to detect an effect size of 0.25 standard deviations in the case of continuous outcomes and of 10 percentage points in the case of dichotomous outcomes. The calculations account for the experimental design (randomization at two levels) and for intra-cluster (intra-CAIF center) correlation of the outcome. The E2P index and subscales do not include the full set of E2P items in the original scale. Basic controls= Strata, child's age and gender, and mother's education. Additional controls= Mother's age, time elapsed since 1st message, negative shocks, other children in household, other adults in household and intact family. Std. Dev.= Standard Deviation. n/a=not available. Robust standard errors in parentheses. * p<0.1, ** p<.05, *** p<.01

Chapter 3

The Effect of Maternal Education on Infant Health: Evidence from an Expansion of Preschool Facilities[1]

3.1 Introduction

It is commonly held that maternal education is an important determinant of children's health. While numerous studies have documented a positive correlation between mother's schooling and child health (see a review in Grossman, 2006), the evidence showing causal effects is scarce. Most studies on the effects of maternal schooling on infant health look at extensions of schooling at the end of the school trajectory and results are mixed.

This chapter studies the effect of an expansion of public preschool facilities in Uruguay on health at birth of the next generation. I use an infrastructure program implemented in the mid 1990's by the Uruguayan government that substantially increased the availability of preschool facilities. The program created approximately 36,000 places between the years 1995 and 2000, which represents an increase in enrollment of 52% (ANEP, 2007). I evaluate health at birth of the offspring of mothers that were exposed to the reform when they were 4 years old. The identification strategy exploits variation induced by the differential timing and intensity of the construction across regions. I assess the robustness of my estimates to the recent developments in the two-way fixed effects regressions' literature using the method proposed by De Chaisemartin and d'Haultfoeuille (2020).

The importance of infant health is widely acknowledged. Infants that are born with low birth weight have worse outcomes both in the short-run and the long-run, including higher mortality within the first year of life and lower educational attain-

[1]This chapter is based on Bloomfield (2020)

ment and earnings in adulthood (Black et al., 2007).[2] In addition, the literature has found that the health income gradient among adults can be explained in part by poor health at infancy and that health at birth can contribute to the intergenerational transmission of poverty (Case et al., 2005).

Maternal education can affect infant health through several direct and indirect channels. Education can affect children's health directly because it increases the ability to acquire and process health information (Grossman, 1972), so more educated mothers are more efficient in the production and allocation of both their own health and the health of their offspring. Indirect effects of education on children's health may work through fertility decisions and assortative mating. Education entails higher earnings and therefore raises a woman's permanent income, influencing her birthing decisions towards fewer children of higher quality (Becker, 1960). In the same line, better educated women match with better educated and higher income husbands (Behrman and Rosenzweig, 2002) and this reinforces the permanent income effect.

Only recently there have been some attempts to establish the causal effects of maternal schooling on infant health.[3] Most of the evidence comes from studies that look at schooling reforms that increased the school leaving age (Güneş, 2015; Chou et al., 2010; Currie and Moretti, 2003; Breierova and Duflo, 2004; Lindeboom et al., 2009; Doyle et al., 2005; Dinçer et al., 2014).[4] Within this set of studies, results are mixed. Some studies find positive impacts on health at birth as measured by very low birth weight (Güneş, 2015), low birth weight and infant mortality (Chou et al., 2010), birth weight and gestational age (Currie and Moretti, 2003), and child mortality (Breierova and Duflo, 2004). Other studies, however, find no effects of maternal education on infant health (Lindeboom et al., 2009, Doyle et al., 2005, Dinçer et al., 2014).

[2]Black et al. (2007) also find that birth weight has an impact on height, the body mass index and the intelligence quotient at age 18.

[3]A large set of studies document a positive correlation between mother's schooling and child health (Grossman, 2006) but this correlation should, however, not be interpreted causally. Selection bias arises, for example, if mothers of better quality tend to have higher education. The correlation will in that case overestimate the true effect of schooling on birth outcomes.

[4]Güneş (2015) uses a change in the compulsory schooling law in Turkey which extended compulsory schooling from five to eight years as an instrument for maternal education. Chou et al. (2010) look at an extension of compulsory education in Taiwan from 6 to 9 years and exploit differential rates in the expansion across regions. Schooling is instrumented using variations across cohorts in new junior high school openings. Currie and Moretti (2003) instrument maternal education with college openings by county in the US at the time when the mother was aged 17. Breierova and Duflo (2004) use a primary school construction program in Indonesia as exogenous variation in schooling to analyze the effect of parental education on child mortality. Lindeboom et al. (2009) exploit a compulsory schooling reform in 1947 in the UK which changed the age of school exit from 14 to 15 years old. Doyle et al. (2005) use another change in the age of school exit in Britain that occurred in the year 1957 and use grand-parental smoking behavior to instrument parental education and income. Dinçer et al. (2014) use a change in the compulsory schooling law in Turkey as an instrument for schooling.

Increases in maternal education at the beginning of the school trajectory can potentially have large effects on infant health. Preschool education is designed to prepare children for school and fosters the development of cognitive and non-cognitive skills. These early inputs may increase the productivity of investments made later.[5] Evaluations of the long-run impacts of some preschool programs have shown positive effects on schooling. Experimental and non-experimental studies of the Perry Preschool Program and Head Start have found positive effects on outcomes such as educational attainment and earnings (Currie, 2001, Heckman et al., 2010, Garces et al., 2002). In Norway, Havnes and Mogstad (2011) find positive effects of a large-scale expansion of subsidized preschool on educational attainment.

The only paper that focuses on the effect of mothers starting school earlier on infant health is McCrary and Royer (2011). The authors use age-at-school-entry policies in California and Texas and exploit the fact that the year in which a person starts school is a discontinuous function of the exact date of birth. The authors find that starting school early has only small effects on infant health and does not affect fertility or prenatal behaviors such as smoking rates and the use of prenatal care.

This study contributes to the literature by providing evidence of the effects of an expansion at an even earlier grade of education than analyzed by McCrary and Royer (2011). I analyze whether a mother's participation in a preschool program during age 4 impacts the health of her offspring. The main database used in the analysis compiles information from vital statistics natality micro-data for the years 2008-2017. This database has information of pregnancy outcomes and parents' characteristics of all registered births in Uruguay. I combine natality data with a measure of availability of preschool places by region and year that I construct using school level data provided by the National Administration of Public Education.

I find that the the expansion of preschool facilities improved health at birth of the next generation. As measures for infant health, I use low birth weight, very low birth weight and extreme low birth weight and indicators for whether the child was born premature, very premature or extremely premature. The incidence of extreme prematurity decreases among first-born children of mothers exposed to the preschool expansion. When exploring potential pathways, I find that mothers that were exposed to the reform have more years of completed education and a larger probability of having more than seven prenatal checkups during pregnancy.

My findings also indicate that the expansion of preschool places had an impact on fertility. I find that age of the mother at birth increases and that the probability of motherhood decreases. Importantly, the latter effect appears mainly because of a reduction in teenage pregnancies. Fertility is an issue of interest in its own right, but

[5] In the economics literature, Cunha and Heckman (2007) make a strong case that early investments benefit from self-productivity and dynamic complementarities.

it can create concerns for the interpretation of health at birth outcomes. Because changes in fertility in the direction observed in this chapter can naturally lead to an improvement in health at birth results, I estimate the sensitivity of my health at birth estimates to changes in fertility. Following Lee (2009) I conduct a bounding analysis in which I assume different health at birth scenarios for missing babies. Overall, I find that the estimated coefficients on health at birth are qualitatively the same as those observed in the main analysis.

The remainder of this chapter is organized as follows. Section 3.2 describes the schooling reform used in the analysis. Sections 3.3 and 3.4 describe the data and the identification strategy, respectively. Results are presented in Section 3.5 and I conclude in Section 3.6.

3.2 The Reform

Uruguay is a small middle-income country. Around half of the country's 3.2 million inhabitants are concentrated in the capital city, Montevideo, and the rest of the population is distributed across 18 other regions. One quarter of the total population are children aged 0-14. In terms of the education system, Uruguay has a long tradition of publicly provided education. Free schooling is provided to children aged 4 and 5 within primary school premises, or to children aged 3, 4 and 5 in separate kindergartens (ANEP-CODICEN, 2000). Children usually attend public preschool centers 4 hours per day (either in the morning or afternoon shift), 5 days a week, and 9 months a year.

By the mid 1990's, the Uruguayan Government decided to implement a reform to alleviate two features of the education system in Uruguay: grade retention and early dropout. The main pillar of the reform was the creation of extra preschool places with the aim of achieving universal preschool for children aged 4 and 5 (ANEP-CODICEN, 2000). The hope was that the reform would increase the number of years of schooling, not including preschool years, and would facilitate children's insertion and transition through the primary school system.

One of the main constraints for the expansion of preschools was the lack of infrastructure. In 1995, the National Administration of Public Education (ANEP) started a large construction program to expand preschool provision in public schools. Between 1995 and 1999, 414 classrooms were added either because they were newly built or because they were made available after being refurnished. New classrooms were added mainly in the 1,144 existing primary schools and were allocated across the different Uruguayan regions based on a specific allocation rule. Priority in the construction was given to: (i) places with strong demographic growth in corresponding age cohorts in the decade prior to the reform, (ii) deprived areas with low physical investment and (iii) bordering regions with Brazil where the cultural identity needed to be strengthened. This framework generated considerable variation in construction intensity and the supply of preschool facilities among regions.

The reform was successful in increasing preschool participation. Enrollment and attendance rates for children aged 4 and 5 increased substantially between the years 1995 and 2000 (ANEP, 2005, ANEP, 2007). The number of children enrolled in public preschools increased from 49,618 to 84,984, a rise of 71%, while enrollment in private preschools remained relatively stable (the number of pupils increased from 19846 to 20806). The attendance rates of children aged 4 and 5 increased from 65% to 82% between 1995 and 2000 and the increase was more pronounced in the group of 4 year olds. Moreover, the expansion was progressive as it attracted students from more disadvantaged backgrounds. By 1991, attendance rates to preschool of children aged 4 were around 20% for the lowest income quintile, while in 2002 this number was in the order of 60%.

3.3 Data

This chapter combines pregnancy and delivery data with school-level data. This section describes the two datasets used and describes the treatment and outcome variables.

3.3.1 Data Sources

Pregnancy and delivery data comes from the vital statistics natality micro-data for the period 2008-2017. This database provides information on all registered live births in Uruguay. Registered births are around 98% of all pregnancies in the country and the dataset covers on average almost 48,000 births per year. Starting from 2008, the vital statistics provide the following information: (i) parents' characteristics such as year and region of birth, years of education and marital status, (ii) number of previous pregnancies of the mother, (iii) prenatal care utilization, and (iv) birth outcomes including birth weight and gestational week in which the birth occurred.

Apart from natality data, I use school-level data from the Monitor Educativo de Enseñanza Primaria, an administrative registry produced by the Department of Research and Statistics of ANEP. This source provides information on preschool and primary education in Uruguay since 1992 for all public schools. The database contains information on each school's: (i) location, (ii) enrollment by level, number of groups and group size, and (iii) student's educational outcomes (insufficient attendance, repetition and dropouts). The administrative registry has a 100% coverage in all years. I consider information for the period 1992-2000, which includes cohorts exposed and not exposed to the reform, to construct a measure of availability of preschool places.

The two data sources are merged and form a pooled cross-section of mothers born between 1988 and 1996 that gave birth in the years 2008-2017. Due to a low number of births for women younger than 15 (less than 1% of the sample), I restrict the sample to mothers older than 14 years of age which leaves 134,140 births of mothers aged 15 to 29 with complete information for the birth outcomes used in the analysis. Table 3.i in the Appendix shows the corresponding age for each pair of birth-cohort and year of observation in the sample.

I also restrict the sample to first-time mothers which leaves a final restricted sample of 66,592 observations. This sample includes only women that are giving birth for the first time and, hence, constitutes a more homogeneous group than the full sample. Moreover, as discussed by McCrary and Royer (2011), the group of first-time mothers is more comparable to other samples of women that have been analyzed in the literature.

3.3.2 Treatment Variable

The school-level data enables the construction of a measure for availability of public preschool places for children aged 4.[6] Availability of preschool places per child is a measure for treatment intensity that varies by department and cohort of birth of the mother. It is constructed by multiplying the total number of groups in the mother's department of birth, in the year in which she was 4 years old by an average of 25 students per group and dividing this number by the population of the corresponding age in that department and year.[7] [8] Even though exposure varies according to the department where the mother lived at the age of 4, department of birth is preferable to assign treatment intensity because it is not subject to endogenous migration.

Table 3.1 shows the availability of preschool places per child by year and department. On average, available preschool places per child were 0.4 for 4-year-olds in the period 1992-2000. The growth of preschool places between 1992 and 2000 averaged 0.3 preschool places per child and it was different across departments. For example, between the years 1992 and 2000, Maldonado increased availability of preschool places per child by 381% (from 0.11 to 0.53), while Rocha increased its availability of preschool places per child by 55% (from 0.34 to 0.53). The availability of preschool

[6] I consider data of schools and kindergartens.

[7] I exclude data from rural schools. These are extremely small schools located in the countryside. On average, rural schools have 4 students in preschool and 22 students in the 6 grades of primary education while other schools have 39 and 188 students respectively. Rural schools represent 45% of schools in my sample, but they cover a very small fraction of students (6% of preschool students).

[8] Population data comes from the Uruguayan Population Projections by year and age provided by the Uruguayan National Institute of Statistics.

places for 5-year-olds also increased in some departments during the reform, but in a considerable smaller magnitude (see Table 3.ii in the Appendix).

Table 3.1: Availability of preschool places per child by year and department for 4-year-olds

Region	Year									Increase
	1992	1993	1994	1995	1996	1997	1998	1999	2000	1992-2000
Montevideo	0.23	0.25	0.25	0.20	0.29	0.35	0.41	0.43	0.43	89%
Artigas	0.15	0.18	0.17	0.13	0.17	0.31	0.34	0.48	0.45	202%
Canelones	0.21	0.20	0.22	0.18	0.23	0.36	0.50	0.46	0.48	125%
Cerro Largo	0.15	0.20	0.15	0.22	0.28	0.26	0.36	0.36	0.40	168%
Colonia	0.32	0.29	0.31	0.31	0.38	0.54	0.48	0.49	0.53	64%
Durazno	0.30	0.25	0.23	0.28	0.38	0.37	0.58	0.50	0.59	93%
Flores	0.39	0.32	0.39	0.39	0.52	0.76	0.63	0.62	0.50	30%
Florida	0.37	0.34	0.39	0.34	0.39	0.46	0.68	0.64	0.65	77%
Lavalleja	0.34	0.45	0.42	0.45	0.48	0.69	0.63	0.70	0.59	75%
Maldonado	0.11	0.11	0.16	0.21	0.19	0.28	0.59	0.51	0.53	381%
Paysandu	0.12	0.14	0.11	0.12	0.14	0.33	0.49	0.44	0.39	219%
Rio Negro	0.30	0.32	0.32	0.32	0.40	0.40	0.44	0.49	0.61	105%
Rivera	0.20	0.21	0.20	0.27	0.39	0.44	0.39	0.53	0.60	203%
Rocha	0.34	0.25	0.32	0.37	0.37	0.49	0.50	0.44	0.53	55%
Salto	0.16	0.20	0.20	0.15	0.13	0.33	0.40	0.43	0.44	179%
San Jose	0.20	0.22	0.24	0.26	0.29	0.57	0.63	0.58	0.63	213%
Soriano	0.29	0.27	0.29	0.29	0.27	0.47	0.37	0.45	0.52	84%
Tacuarembo	0.42	0.40	0.38	0.31	0.43	0.57	0.59	0.65	0.68	63%
Treinta y Tres	0.24	0.21	0.32	0.38	0.35	0.48	0.47	0.44	0.60	155%

Note: Availability of preschool places per child of age 4 is calculated as the number of groups opened for 4-year olds by region and year multiplied by an average of 25 students per group and divided by the number of children aged 4 in each region in the corresponding year (obtained from the Uruguayan National Institute of Statistics).

3.3.3 Outcome Variables

In Table 3.2 I define the outcome variables used in the analysis. Birth outcomes are the main dependent variables. To measure health at birth, I focus on low birth weight, a measure that generally is considered as an indicator for intrauterine growth retardation during pregnancy and/or being born premature. In particular, I use indicators for low birth weight, very low birth weight, extreme low birth weight, premature, very premature, and extremely premature.[9] I define the latter thresholds according to the definitions listed in the International Statistical Classification

[9]Severe cases of low birth weight such as very low birth weight and extreme low birth weight are linked to higher risk of death during the newborn period, poor school performance and adverse outcomes during adulthood (Hack et al., 2002, 1994; Hack and Fanaroff, 1999).

of Diseases and Related Health Problems (ICD-10) codes of the World Health Organization. Children born with extreme low birth weight or extreme prematurity have a higher risk of facing health difficulties later on, so it is worthwhile exploring whether results are sensitive to these margins. I also consider outcomes that could shed light on potential channels for changes in birth outcomes. These variables relate to maternal and paternal characteristics, maternal health behavior, and fertility decisions.

Table 3.2: Description of variables

Variable	Definition
Birth outcomes	
Birth weight	Weight in grams.
Low birth weight indicators	Binary variables. Each variable equals 1 if birth weight is below threshold, 0 otherwise. The thresholds considered in the analysis are: 2500g, 1500g, and 1000g. These thresholds are referred to as: low birth weight, very low birth weight and extreme low birth weight, respectively.
Gestational weeks	Weeks of gestation at birth.
Prematurity indicators	Binary variables. Each variable equals 1 if birth occurred before the threshold week of gestation, 0 otherwise. The thresholds considered in the analysis are: 37 weeks, 32 weeks, 28 weeks. These thresholds are referred to as: premature, very premature and extremely premature, respectively.
Parental characteristics	
Mother's years of education	Completed years of education of the mother.
Father's years of education	Completed years of education of the father.
Mother's age at birth	Age in years at first born's birth.
Mother and father of child live together	Binary variable. 1 if father and mother of the child live together at child's birth, 0 otherwise.
Maternal health behavior	
Prenatal care in first trimester of pregnancy	Binary variable. Equals 1 if woman initiated prenatal care during her first trimester of pregnancy, 0 otherwise.
Mother had more than 7 prenatal checkups during pregnancy	Binary variable. Equals 1 if woman had more than 7 prenatal checkups during pregnancy, 0 otherwise.
Fertility outcomes	
Total fertility	Number of first-borns by each cohort and region until 2015 divided by the population of the corresponding cohort. This measure is aggregated at the region-cohort level.
Fertility by age	Number of first-borns by age, cohort and region until 2015 divided by the population of the corresponding cohort. This measure is aggregated at the region-cohort-age level.

Note: Sources for construction of variables are the Uruguayan natality vital statistics and population projections by year and age provided by the Uruguayan National Institute of Statistics.

Table 3.3 shows sample statistics for birth outcomes, maternal and paternal characteristics, and maternal health behavior. The average weight of newborns is 3226 grams and the incidence of low birth weight for the thresholds of 2500, 1500 and 1000 grams is 8%, 1.2% and 0.5% respectively. The likelihood that the baby is born before the 37th, 32nd and 28th week of gestation is 8.8%, 1.3% and 0.5%

respectively. The average number of gestational weeks is approximately 38.6 weeks. In terms of maternal and paternal characteristics, both mothers and fathers on average have completed 9 years of education completion. In this sample, females are aged 21 years on average when giving birth for the first time. This value is lower than the population average age of first motherhood in the period 2008-2017 which is 24. Slightly more than half of the mothers report living with the father of the child, which is in accordance with population statistics.[10] In terms of prenatal care, 69% of mothers have a prenatal control in the first trimester of pregnancy. On average, 78% of first-time mothers visit the doctor more than 7 times during pregnancy.

Table 3.3: Descriptive statistics

	Mean	s.d.	N
Birth outcomes			
Weight (in grams)	3226.138	553.851	66592
Low birth weight (<2500g)	0.079	0.270	66592
Very low birth weight (<1500g)	0.012	0.110	66592
Extreme low birth weight (<1000g)	0.005	0.068	66592
Premature (<37 weeks of gestation)	0.088	0.283	66592
Very Premature (<32 weeks of gestation)	0.013	0.115	66592
Extreme Prematurity (<28 weeks of gestation)	0.005	0.068	66592
Gestational weeks	38.570	1.952	66592
Parental characteristics			
Mother's education in years	9.266	2.607	66592
Father's education in years	8.870	2.586	37794
Average years of education between mother and father	9.314	2.238	37794
Mother's age at birth	20.557	3.003	66592
Mother and father of child live together	0.544	0.498	66592
Maternal health behavior			
Prenatal care in first trimester of pregnancy	0.693	0.461	66592
More than 7 prenatal checkups during pregnancy	0.781	0.413	66592

Note: s.d.=standard deviation, N=number of observations. Mother completed primary school=1, mother and father of child live together=1, prenatal care in first trimester=1 if mother had at least one prenatal control during the first trimester of pregnancy, more than 7 prenatal checkups=1.

[10]According to the Uruguayan National Survey of Nutrition, Child development and Health of 2018, the percentage of first-time mothers younger than 28 years that live with the father of the child during the first year after giving birth is 53%.

3.4 Empirical Strategy

The aim of this chapter is to estimate the effect of the expansion of preschool places on the health at birth of children of exposed mothers. Following Duflo (2001) and Berlinski and Galiani (2007) the empirical strategy relies on a generalized difference-in-differences strategy that combines differences across regions in the number of facilities built with differences in exposure across cohorts induced by the timing of the program. In my estimations, I control for region and cohort fixed effects: region fixed effects control for constant characteristics at the region level that are fixed over time and cohort fixed effects control for unobserved differences across cohorts. I also include year of child's birth fixed effects to control for common effects shared by mothers giving birth at the same time such as the economic conditions or particular policies.

I evaluate the impact of the preschool expansion on health at birth outcomes, maternal and paternal characteristics and prenatal care by estimating equations of the following form using Ordinary Least Squares (OLS):

$$Y_{icdt} = \alpha_1 + \delta_{1d} + \gamma_{1c} + \rho_{1t} + \beta_1 Stock_{cd} + \varepsilon_{icdt} \tag{3.1}$$

where Y_{icdt} is the outcome of interest for the birth of child i, whose mother was born in cohort c and region d, and is observed in year t; δ_{1d} are region fixed effects; γ_{1c} are cohort fixed effects and ρ_{1t} are year of child's birth fixed effects; and $Stock_{cd}$ is a measure for the availability of preschool places per child. β_1 captures the average effect of an extra place available per child on the outcome variable of interest. I adjust the standard errors for clustering at the region times cohort level.

The reform may have influenced decisions such as the number of children and the timing of childbearing. To study the impact of the reform on fertility I estimate the impact of the expansion of preschool places per child on the overall probability of motherhood and on the probability of motherhood by age using the following equation:

$$F_{cd} = \alpha_2 + \delta_{2d} + \gamma_{2c} + \beta_2 Stock_{cd} + \varepsilon_{cd} \tag{3.2}$$

where F_{cd} corresponds to total fertility or fertility by age. Total fertility is defined as the number of first-borns by cohort and region until 2015 divided by the population of the corresponding cohort. This measure is aggregated at the region-cohort level. Fertility by age is the number of first-borns by age, cohort and region until 2015 divided by the population of the corresponding cohort. β_2 captures the average effect of an extra place available per child on the outcome variable. Standard errors (ε_{cd}) are clustered at the region times cohort level. I also adjust standard errors to account for multiple hypotheses using the Romano-

Wolf correction described in Romano and Wolf (2005a,b, 2016) and report adjusted p-values.[11]

The identification strategy relies on the parallel trends assumption, which implies that the trend in the outcome variable should not have been systematically different in regions where the program constructed more preschool places and regions where the program constructed fewer preschool places prior to the expansion policy. In other words, the main identification assumption is that in the absence of an increase in the availability of preschool places, changes in health at birth would not have been systematically different between mothers from regions/cohorts with low versus high exposure. To verify the robustness of my estimates, I will conduct several checks that I describe below.

The treatment intensity of implementation of the policy could be correlated with specific pre-treatment trends between regions. Indeed, treatment intensity was based on the demographic growth preceding the implementation of the construction program. Therefore, as a first robustness check, I include differential trends by region in my estimations. Equation 3.1 transforms into:

$$Y_{icdt} = \alpha_3 + \delta_{3d} + \gamma_{3c} + \rho_{3t} + \beta_3 Stock_{cd} + \sum_{d=1}^{19} (\mathbb{1}(D = d) * t_c)\phi_{3d} + e_{icdt} \qquad (3.3)$$

where $\mathbb{1}(D = d) * t_c$ represents the interaction of a dummy for region d and a cohort indicator (t_c) that takes values from 1 (for year 2008) to 6 (for year 2013).

The common trend assumption could also be violated if changes in health at birth would have happened faster in the absence of the program in regions where the starting enrollment rates in preschool were higher and if the allocation of preschool places was correlated to the starting enrollment rates. To address this concern, following Duflo (2001), I control for possible omitted time-varying region-level factors that may be correlated with pre-program enrollment rates by adding to the main model the interaction of available preschool places per child by region in 1995 with fixed effects by cohort. Equation 3.1 is adjusted in the following way:

$$Y_{icdt} = \alpha_4 + \delta_{4d} + \gamma_{4c} + \rho_{4t} + \beta_4 Stock_{cd} + \sum_{c=1}^{9} (Stock_{1995d} * d_c)\phi_{4c} + \mu_{icdt} \qquad (3.4)$$

where $Stock_{1995d} * d_c$ represents the interaction of the stock of available preschool places in 1995 in each region ($Stock_{1995d}$) and cohort dummies (d_c).

Finally, I perform additional placebo regressions to verify the robustness of my estimates. I estimate Equation 3.1 using pre-treatment cohorts. If trends between

[11] I consider two families of outcomes: one that includes the outcome birth weight and the low birth weight indicators and the other that includes the prematurity indicators.

regions with different treatment intensity are the same in the pre-treatment period, the expectation is that there should be no impact of the expansion of preschool places on health at birth on those years.

A recent literature has questioned the validity of regressions using two-way fixed effects. These papers point out that in such identification strategy, if only common trends are assumed, the estimated treated effect is a weighted sum of the effect of treatment in each group and time period. Some of the weights can be negative and this may bias or even change the sign of the true average treatment effect. De Chaisemartin and d'Haultfoeuille (2020), propose computing the DID_M estimator, that only relies on the common trends assumption and is valid even if the treatment effect is heterogeneous over time and between groups. The DID_M estimator can be computed using the *didmultiple_gt* Stata package (de Chaisemartin et al., 2019). In Section 3.5.4 I analyze the sensitivity of my estimates to this adjustment.

3.5 Results

In this section I present the results of the analysis. First, I provide evidence of the effects of the reform on health at birth. Second, I report the effects of the expansion of preschool places on potential pathways such as parents' characteristics, prenatal care and fertility. Third, I study heterogeneous effects. Finally, I analyze the robustness of my estimates.

3.5.1 Effects of the Expansion of Preschool Places on Birth Outcomes

Table 3.4 shows the results of estimating Equation 3.1 (Column (1)), Equation 3.3 (Column (2)), and 3.4 (Column (3)) for birth outcomes. The expansion of preschool places implied an improvement in health at birth as measured by extreme prematurity. For every preschool place opened per child, the probability of giving birth to a child before week 28 of gestation decreased by 1.3 percentage points (see Column (1)). This effect maintains significance at the 1% level when considering standard errors that are adjusted for multiple hypothesis testing,[12] when controlling for differential trends by region and when controlling for possible omitted time-varying region-level factors that may be correlated with pre-program enrollment rates. As the average increase in preschool places between the years 1992 and 2000 is 0.3, the magnitude of the coefficient can be interpreted as follows: an increase of 0.3 preschool places per child leads to a decrease in extreme prematurity of 0.4

[12]The effect also maintains significance at the 1% level when grouping all outcomes together in one family.

(0.013 x 0.3) percentage points. There is also a significant reduction in the likelihood of low birth weight in the range of 4.5 percentage points when estimating Equations 3.1 and 3.4, and a significant reduction in the likelihood of extreme low birth weight in the range of 1 percentage points when estimating Equation 3.3, but these effects become insignificant when estimating Equation 3.3, and 3.1 and 3.4, respectively.

Table 3.4: Effects of the expansion of preschool places per child on birth outcomes

Dependent variable	(1)	(2)	(3)
Birth weight (in grams)	-14.249	-27.723	-10.283
	(46.742)	(53.966)	(46.727)
	[0.911]		
Low birth weight (<2500g)	-0.043*	-0.034	-0.052**
	(0.025)	(0.031)	(0.025)
	[0.089]		
Very low birth weight (<1500g)	-0.003	-0.009	-0.006
	(0.011)	(0.014)	(0.011)
	[0.910]		
Extreme low birth weight (<1000g)	-0.008	-0.012*	-0.007
	(0.006)	(0.007)	(0.006)
	[0.158]		
Premature (<37 weeks)	-0.019	-0.032	-0.019
	(0.025)	(0.032)	(0.025)
	[0.604]		
Very premature (<32 weeks)	-0.005	-0.018	-0.006
	(0.012)	(0.014)	(0.011)
	[0.604]		
Extreme prematurity (<28 weeks)	-0.013***	-0.018***	-0.012***
	(0.005)	(0.006)	(0.005)
	[0.010]		
Interaction of region fixed effects and cohort indicator	No	Yes	No
Interaction of stock in 1995 and fixed effects by cohort	No	No	Yes

Note: Table reports results for the estimation of Equation 3.1 (Column (1)), Equation 3.3 (Column (2)), and 3.4 (Column (3)) for several dependent variables using OLS. Estimations include cohort fixed effects, region fixed effects and year fixed effects. Standard errors, reported in parentheses, are clustered at the region times cohort level. I report p-values that are adjusted for multiple hypothesis testing in squared brackets. Number of observations is 66592. * p<.1, ** p<.05, *** p<.01.

3.5.2 Mechanisms

Maternal and paternal characteristics and prenatal care

Table 3.5 shows the results of estimating the effect of the preschool expansion on parental characteristics as well as on prenatal care. Mother's education increased by 0.6 years (see Column (1)) for every preschool place opened per child.[13] The

[13]Berlinski et al. (2008) exploit the same expansion of preschool places in Uruguay using a within-household estimator. The authors find that by the age of 15 treated children accumulated 0.8 extra years of education in comparison to untreated siblings and that this works through a fall

effect maintains significance at 1% when controlling for the interaction of the stock of preschool places in 1995 with fixed effects by cohort and at 10% when controlling for the interaction of region fixed effects with a cohort indicator. Moreover, the expansion of preschool increased the age at motherhood by 0.13 years for every preschool place that was opened per child. In terms of father's characteristics, there is no effect of the reform on the number of completed years of education, and the effect of the reform on the probability that mother and father live together is not robust to the different specifications considered in the analysis. Regarding prenatal care, the likelihood that the mother had more than 7 prenatal checkups increased by 10 percentage points per preschool place opened per child (or by 3 percentage points - 0.10 x 0.3 - for every 0.3 preschool places opened per child). These effects are robust to the specifications detailed in Equation 3.3 and Equation 3.4.

Table 3.5: Effects of the expansion of preschool places per child on maternal and parental characteristics and prenatal care

Dependent variable	(1)	(2)	(3)
Mother's years of education	0.613***	0.379*	0.597***
	(0.182)	(0.196)	(0.191)
Father's years of education	0.370	-0.107	0.326
	(0.253)	(0.255)	(0.252)
Average years of education of mother and father	0.365*	0.036	0.272
	(0.203)	(0.213)	(0.207)
Mother and father live together	0.094*	-0.058	0.153***
	(0.051)	(0.043)	(0.050)
Age of the mother at birth	0.126***	0.091**	0.139***
	(0.040)	(0.037)	(0.042)
More than 7 prenatal checkups during pregnancy	0.101***	0.075*	0.114***
	(0.033)	(0.040)	(0.034)
Care in first trimester	0.051	0.013	0.058
	(0.044)	(0.050)	(0.046)
Interaction of region fixed effects and cohort indicator	No	Yes	No
Interaction of stock in 1995 and fixed effects by cohort	No	No	Yes

Note: Table reports results for the estimation of Equation 3.1 (Column (1)), Equation 3.3 (Column (2)), and 3.4 (Column (3)) for several dependent variables using OLS. Estimations include cohort fixed effects, region fixed effects and year fixed effects. Standard errors, reported in parentheses, are clustered at the region and cohort level. Number of observations is 66592 for all outcomes except for father's years of education which is 37794. * p<.1, ** p<.05, *** p<.01.

Fertility

Table 3.6 shows the effect of the expansion of preschool places per child on the probability of motherhood. I estimate Equation 3.2 and report the pooled coefficient for

in grade retention rates in the school trajectory and a reduction in dropout rates. The effect of the expansion of preschool places at age 15 seems larger than the one found in this chapter, suggesting that the gap between treated and untreated children widens during adolescence.

each age and for the overall sample. The number of observations in each regression are reported in the last column. For example, 171 observations were used in the regression that considers fertility of women aged 19 as the outcome variable. There is one observation for each combination of the 19 regions and 9 cohorts. I also show the mean number of women that gave birth to their first child at each age. For example, 2.4% of all women born between 1988 and 1996 that gave birth to their first child between 2008 and 2017, were aged 27 when giving birth (see the third column under age 27).[14] The estimate of the reform on the probability of motherhood when considering all ages is negative and significant at the 5% level. For every preschool place opened per child, overall fertility rate decreases by 11 percentage points. The effect essentially comes from a reduction in teenage births (at ages 16 and 17) and at age 28. Indeed, when I estimate the impact of the reform on overall fertility without considering births of females aged 16, 17 and 28, I find that the coefficient is insignificant.

[14]Note that only those born in 1988-1990 could have given birth at age 27 between 2008 and 2017 (see Table 3.i in the Appendix). Women born in later cohorts are observed at younger ages in my sample.

Table 3.6: Effect of the expansion of preschool places per child on probability of motherhood

Age	(1)	(2)	(3)	(4)
Panel A: Fertility by age				
15	-0.008	0.012	0.011	95
16	-0.039***	0.013	0.027	114
17	-0.037**	0.017	0.034	133
18	-0.023	0.016	0.040	152
19	-0.022	0.018	0.042	171
20	-0.019	0.014	0.042	171
21	-0.009	0.014	0.038	171
22	-0.004	0.012	0.034	152
23	-0.000	0.014	0.031	133
24	0.016	0.014	0.026	114
25	-0.008	0.027	0.026	95
26	0.025	0.027	0.025	76
27	0.051	0.035	0.024	57
28	-0.104**	0.038	0.020	38
29	0.007	0.017	0.011	19
Panel B: Overall fertility				
All ages	-0.110**	0.054	0.320	171
All ages except age 16, 17 and 28	-0.026	0.047	0.271	171
Panel C: Fertility when including additional births				
Age 16 including added births	-0.027	0.016	0.031	114
Age 17 including added births	-0.031	0.019	0.038	133
Age 28 including added births	0.136	0.166	0.032	38
All ages including added births	-0.029	0.060	0.328	171

Note: Table shows results of the OLS estimation of the effect of the expansion of preschool places per child on overall fertility and on the probability of motherhood by age group. Overall fertility and probability of motherhood by age group are defined in Table 3.2. Column (1) reports the coefficient that corresponds to the independent variable available preschool places per child. Estimations include cohort fixed effects and region fixed effects. Standard errors, reported in Column (2), are clustered at the region times cohort level. Sample means are reported in Column (3). Number of observations included in the estimations are reported in Column (4). * p<.1, ** p<.05, *** p<.01.

Changes in fertility decisions could be a potential channel by which education can affect health at birth. In order to analyze the extent to which changes in fertility explain the findings reported in Table 3.4, I conduct two analyses. First, I estimate the impact of the reform on health at birth of the next generation excluding mothers that gave birth at ages 16, 17 and 28. Second, I conduct a bounding analysis. Overall, the takeaway from these analyses is that fertility is not a strong pathway underlying the health at birth estimates. A detailed description of both analyses can be found below.

When considering first-time mothers of all ages excluding those aged 16, 17 and 28, the impact of the reform on fertility is non-significant (see Table 3.6). In this sample we should not expect that fertility is a mechanism behind the observed effects on health at birth. In Table 3.7 I estimate the effect of the reform on health at birth of first-borns of mothers that are not aged 16, 17 and 28. The results are qualitatively very similar to those found when including mothers of all ages in the estimation. Hence, my interpretation is that fertility does not play a significant role when considering the whole age distribution.

Table 3.7: Effects of the expansion of preschool places per child on birth outcomes excluding mothers aged 16, 17 and 28

	Coefficient	s.e.
Birth weight (in grams)	33.506	(53.912)
Low birth weight (<2500g)	-0.051*	(0.029)
Very low birth weight (<1500g)	-0.009	(0.012)
Extreme low birth weight (<1000g)	-0.010*	(0.006)
Premature (<37 weeks)	-0.043	(0.030)
Very premature (<32 weeks)	-0.008	(0.012)
Extreme prematurity (<28 weeks)	-0.011**	(0.005)

Note: Table reports results for the estimation of Equation 3.1 for several dependent variables using OLS. Estimations include cohort fixed effects, region fixed effects and year fixed effects. Standard errors, reported in parentheses, are clustered at the region times cohort level. Number of observations is 66592. * p<.1, ** p<.05, *** p<.01.

The second approach to analyze the sensitivity of my estimates to changes in fertility is a bounding analysis. Given the negative impact of the reform on fertility, in the absence of the reform, I would expect to have observed more births of first-time mothers.[15] To bound the effect of the reform on birth outcomes, I add more

[15] If the births that are not observed in my sample were the more healthy ones then I would expect that, in the absence of the reform, the improvement on health at birth would have been even larger than the one reported in Table 3.4. In that case, my estimates of the effect of the reform on low birth weight and prematurity are a lower bound (in absolute terms) of the true

children to the regions where the expansion was largest. In particular, I include births to regions that more than doubled their availability of preschool places per child between 1992 and 2000. In Table 3.iii I report the number of children that I add to each of these regions.[16] When estimating the effect of the reform on the age-specific fertility rate for ages 16, 17 and 28 including added births, the effect is no longer significant (see Table 3.6). In addition, the effect on the overall fertility rate when including the additional children is no longer significant either.

Next, I assume different birth-weight-scenarios for the observations that were added. In particular, I assume the following five situations regarding the health at birth of the children that I incorporate: (1) birth weight is 900 grams and gestational length is 27 weeks, (2) birth weight is 1400 grams and gestational length is 31 weeks, (3) birth weight is 2400 grams and gestational length is 36.5 weeks, (4) birth weight is 3200 grams and gestational length is 39 weeks, and (5) birth weight is 4000 grams and gestational length is 40 weeks. In this sense, scenario (1) is an extreme case in which the unobserved births are very unhealthy while scenario (5) assumes the opposite.

Table 3.8 shows the results of estimating Equation 3.1 using the augmented sample of first-time mothers for the different birth-weight-scenarios of the missing observations. As outcome variables, I consider health at birth indicators. In Scenario 4 and Scenario 5, the results are qualitatively the same than those reported in Table 3.4: the incidence of low birth weight and extreme prematurity decreases among mothers that were exposed to the reform. In Scenarios 2 and 3, the effects of the reform on extreme prematurity remains but the effect on low birth weight disappears. When including observations of extremely unhealthy babies (Scenario 1), all the effects observed in Table 3.4 disappear. There are some significant effects but these have the opposite sign than expected. Considering that the mean birth weight among mothers aged 16, 17 and 28 is 3173 grams, my preferred Scenario is number 4. In light of the evidence of Table 3.8 and the analyses shown above, my conclusion is that fertility is not a strong pathway underlying the health at birth estimates.

effect. If, on the contrary, the births that are not observed in my sample were the more unhealthy births, then, in the absence of the reform, I would expect to find a smaller improvement in health at birth than the one found.

[16] I assume that the births that I include belong to mothers that where born in 1996 and that gave birth in 2015.

Table 3.8: Effects of the expansion of preschool places per child on birth outcomes considering different scenarios for missing observations

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
	900 grams 27 gestational weeks	1400 grams 31 gestational weeks	2400 grams 36.5 gestational weeks	3200 grams 39 gestational weeks	4000 grams 40 gestational weeks
Birth weight (in grams)	-388.369*	-306.675*	-143.288*	-12.578	118.131
	(216.459)	(170.991)	(84.463)	(45.332)	(90.441)
Low birth weight (<2500g)	0.108	0.108	0.108	-0.055**	-0.055**
	(0.091)	(0.091)	(0.091)	(0.025)	(0.025)
Very low birth weight (<1500g)	0.159*	0.159*	-0.004	-0.004	-0.004
	(0.093)	(0.093)	(0.011)	(0.011)	(0.011)
Extreme low birth weight (<1000g)	0.155*	-0.008	-0.008	-0.008	-0.008
	(0.092)	(0.006)	(0.006)	(0.006)	(0.006)
Premature (<37 weeks)	0.129	0.129	0.129	-0.035	-0.035
	(0.087)	(0.087)	(0.087)	(0.026)	(0.026)
Very premature (<32 weeks)	0.156*	0.156*	-0.007	-0.007	-0.007
	(0.092)	(0.092)	(0.012)	(0.012)	(0.012)
Extreme prematurity (<28 weeks)	0.150	-0.013***	-0.013***	-0.013***	-0.013***
	(0.093)	(0.005)	(0.005)	(0.005)	(0.005)

Note: Standard errors, reported in parentheses, are clustered at the region times cohort level. * p<.1, ** p<.05, *** p<.01.

3.5.3 Heterogeneous Impacts

Given that the expansion was not targeted to a specific group of the population, in this subsection I explore whether there are differential effects in some groups. I study whether the effects on birth outcomes are specific to a certain group of women or are generalizable. More specifically, I focus on whether the results vary according to the socioeconomic level of the region of birth of the mother in the period previous to the reform. This analysis sheds light on whether the reform benefited disadvantaged regions more or less than other regions. As a proxy for the disadvantagedness of regions, I use the unemployment rate. I split the sample into regions with higher and lower than the median unemployment rate in the period 1992-1994.

Table 3.9 reports results from estimating heterogeneous impacts by unemployment level in the region of birth of the mother in the period previous to the reform. I find that there is a statistically significant difference in the likelihood of extreme low birth weight between mothers that were born in regions with different socioeconomic levels. The effect of the reform on extreme low birth weight was larger in absolute terms among mothers born in regions with high unemployment. For the other outcomes I cannot reject that the effects are the same across the socioeconomic level of the region of birth of the mother, however, heterogeneous effects cannot be ruled out entirely due to a low statistical power to detect differential effects.

Table 3.9: Heterogeneous impacts by unemployment level in the department of birth of the mother in the period previous to the reform

	Low unemployment		High unemployment		Difference	
	Coefficient	s.e.	Coefficient	s.e.	Difference	s.e.
	(1)	(2)	(3)	(4)	(5)	(6)
Birth weight (in grams)	-75.416	(101.465)	38.845	(48.908)	114.261	(112.016)
Low birth weight (<2500g)	-0.038	(0.054)	-0.043	(0.030)	-0.005	(0.062)
Very low birth weight (<1500g)	0.003	(0.018)	-0.003	(0.015)	-0.006	(0.023)
Extreme low birth weight (<1000g)	0.011	(0.010)	-0.015**	(0.007)	-0.026**	(0.012)
Premature (<37 weeks)	-0.028	(0.046)	-0.032	(0.030)	-0.004	(0.055)
Very Premature (<32 weeks)	-0.021	(0.020)	-0.012	(0.015)	0.009	(0.025)
Extreme prematurity (<28 weeks)	-0.017**	(0.007)	-0.012*	(0.007)	0.005	(0.010)

Note: Table reports results of estimating Equation 3.1 for several dependent variables for groups of mothers born in higher or lower than the median unemployment regions. Estimations use OLS. Estimations include cohort fixed effects, region fixed effects and year fixed effects. Standard errors, reported in parentheses, are clustered at the region times cohort level. * p<.1, ** p<.05, *** p<.01.

3.5.4 Robustness

In the main analysis I have controlled for cohort, region and year-of-birth fixed effects, and I have also considered a specification that includes differential trends by region and a specification that controls for the interaction of the stock of preschool places in 1995 with fixed effects by cohort. In this subsection, I test the robustness of my estimates to alternative specifications.

Pre-treatment Cohorts

As a further robustness check, I restrict the sample to those cohorts that were not exposed to the treatment. The assumption of the identification strategy is that, in the absence of treatment, the trends in outcomes would be parallel between regions that were intensively treated and regions that were less treated. In pre-treatment cohorts, we should expect the common trend assumption to hold. If trends between regions with different treatment intensity are the same in pre-treatment cohorts, there should be no association of the expansion of preschool places with the health at birth of children born to women who were 4 years old before the reform. In this subsection, I estimate Equation 3.1 using only pre-treatment cohorts. The sample contains females born in cohorts 1988-1990. These women were aged 17 to 29 when observed giving birth. Those born in 1988 were 7 years old when the expansion takes place.

In Table 3.10 I show that the impact of the reform on prematurity outcomes is insignificant and that I cannot reject equality from zero. For birth weight outcomes I find a pair of significant effects but coefficients have the opposite sign than expected.

Table 3.10: Effects of the expansion of available preschool places for children per child on pre-treatment cohorts

Dependent variable	Coefficient	s.e.
Birth weight (in grams)	-405.592***	(139.808)
Low birth weight (<2500g)	0.022	(0.077)
Very low birth weight (<1500g)	0.101*	(0.052)
Extreme low birth weight (<1000g)	-0.014	(0.023)
Premature (<37 weeks)	0.128	(0.101)
Very premature (<32 weeks)	0.040	(0.048)
Extreme prematurity (<28 weeks)	0.006	(0.017)

Note: Table reports results of estimating Equation 3.1 for several dependent variables using only pre-treatment cohorts. Estimations include cohort fixed effects, region fixed effects and year fixed effects. Standard errors, reported in parentheses, are clustered at the region times cohort level. Number of observations is 18429. * p<.1, ** p<.05, *** p<.01.

DID_M estimator

Two-way fixed effects has been widely-used as econometric method in recent years. De Chaisemartin and d'Haultfoeuille (2020) conduct a survey of all empirical papers between 2010 and 2012 in the American Economic Review and find that 20% of them use two-way fixed regressions to measure the effect of a treatment on an outcome.[17] In two-way fixed effects regressions one would be typically regressing an outcome that is aggregated at the group and time level on group fixed effects, time fixed effects and an independent variable that is aggregated at the group and time level.

A recent body of work questions the validity of the two-way fixed effects estimator in the presence of heterogeneous treatment effects. De Chaisemartin and d'Haultfoeuille (2020) show that if treatment effects are not constant between units or over time then two-way fixed effects regressions identify the expectation of a weighted sum of the treatment effects in every group and every time period. Some of the weights are strictly negative and this may lead to a strictly negative estimated effect even if the treatment effect is strictly positive in every group and at every time period. When the treatment is not binary, negative weights arise from the fact that the identification strategy compares the outcome evolution in groups whose treatment increases more and in groups where treatment increases less (De Chaisemartin and d'Haultfoeuille, 2018).

De Chaisemartin and d'Haultfoeuille (2020) propose an alternative estimator DID_M whose building blocks are difference in differences that compare the outcome evolution in groups going from untreated to treated in both dates (switchers in) and groups untreated at both dates (never treated) and differences in differences comparing the outcome evolution in groups going from treated to untreated in both dates (switchers out) and groups that are treated at both dates (always treated). Therefore DID_M compares switchers to non-switchers making sure that the controls used for a switcher have the same treatment as a switcher in $t-1$. This ensures that the estimator only relies on parallel trends rather than homogeneous treatment effects. The estimator identifies the treatment effect of the switchers at the time they switch.

To analyze the sensitivity of my estimates to the adjustment proposed in De Chaisemartin and d'Haultfoeuille (2020) I compute the DID_M estimator using the *didmultiple_gt* Stata package (de Chaisemartin et al., 2019). In the context of this paper, the DID_M is a weighted average of diff-in-diff estimators comparing the evolution of health at birth between cohort $c-1$ to cohort c, in regions whose treatment changes from $stock_{cd}$ to some other value from $c-1$ to c, and in regions whose treatment is equal

[17] De Chaisemartin and D'Haultfoeuille (2022) update this survey using a slightly different survey method and find that 26% of the most cited American Economic Review papers between 2015 and 2019 use two-way fixed effects regressions.

to $stock_{cd}$ for both cohorts. The estimator uses regions whose treatment does not change between consecutive cohorts as controls.

Given that my treatment is continuous, it is not possible to find any pair of consecutive cohorts between which the treatment of at least one region remains perfectly stable. de Chaisemartin et al. (2019) propose to specify a threshold of stable treatment so that the *didmultiple_gt* Stata command can use that value to determine which regions are used as controls. Moreover, when the treatment takes many values, de Chaisemartin et al. (2019) propose to bin some values of the treatment together when determining the groups whose outcome evolution are compared.

Table 3.11 shows the effects of the expansion of preschool places per child using the DID_M estimator. My preferred threshold of stable treatment is 0.09 which corresponds to the average increase in the availability of preschool places between one cohort and the next in regions that increased the stock of preschool places for children aged 4. I also consider other thresholds for robustness. The treatment variable is grouped in quartiles. By and large, the results of the DID_M estimator qualitatively coincide with those of the main analysis of this paper. Consistently with the fact that two-way fixed effects regressions bias coefficients towards zero, the magnitude of the effects in Table 3.11 is slightly larger. The effect on extreme prematurity using the DID_M estimator is in the range of -0.3 while the one found with a two-way fixed effects regression was -0.2. When using the methodology of de Chaisemartin et al. (2019), I find that for every 0.3 preschool places opened per child extreme prematurity decreases by 0.8 (0.028 x 0.3) percentage points.

Table 3.11: Effects of the expansion of preschool places per child using the DID_M estimator

	Threshold of stable treatment				
	0.080	0.085	0.09	0.095	0.10
	(1)	(2)	(3)	(4)	(5)
Birth weight	-35.368	-104.615	-94.380	-68.874	-150.635
	(110.333)	(102.768)	(61.047)	(66.709)	(101.628)
Low birth weight	-0.119**	-0.038	-0.042	-0.063	0.020
	(0.059)	(0.049)	(0.049)	(0.050)	(0.075)
Very low birth weight	-0.009	0.002	-0.001	-0.011	-0.021
	(0.048)	(0.029)	(0.027)	(0.037)	(0.030)
Extreme low birth weight	-0.011	-0.011	-0.013	-0.018**	-0.025
	(0.023)	(0.015)	(0.023)	(0.008)	(0.020)
Premature (<37 weeks)	0.036	-0.012	-0.007	-0.022	0.006
	(0.129)	(0.053)	(0.044)	(0.053)	(0.040)
Very premature (<32 weeks)	-0.035	-0.021	-0.025	-0.033	-0.039
	(0.042)	(0.039)	(0.031)	(0.027)	(0.034)
Extreme prematurity (<28 weeks)	-0.023**	-0.024	-0.028**	-0.032**	-0.039**
	(0.011)	(0.016)	(0.012)	(0.012)	(0.019)

Note: Table shows results of the expansion of preschool places on birth outcomes using the DID_M estimator as suggested in De Chaisemartin and d'Haultfoeuille (2020). To compute the DID_M estimator, the treatment variable (availablility of preschool places per child aged 4) is grouped into quartiles. The preferred threshold of stable treatment is 0.09 (Column (3)), but I also report results for other thresholds in Columns (1)-(2) and (4)-(5) for robustness.

3.6 Conclusion

This chapter presents estimates of the effects of additional schooling at the beginning of the school trajectory on health at birth of the next generation. I exploit a schooling reform that involved a large construction of preschool places in Uruguay and that occurred at differential rates by region and time.

Using data of availability of preschool places between 1992 and 2000 and birth outcomes in the period 2008-2017 I find an improvement in health at birth of the offspring of those women that were more exposed to the schooling reform. The results suggest a reduction in extreme prematurity for first-time mothers. My estimates are robust to several checks, including the latest advances in the two-way fixed effects methodology.

The findings highlight the importance of education at early years as they show that preschool education has long lasting benefits that can be transmitted across generations. In addition, the evidence in this chapter points to preschool education as a way to reduce the intergenerational transmission of poverty due to poor health at birth.

Potential channels of the observed effects is that exposed mothers in my sample have more years of completed education and are more likely to have more than seven prenatal checkups during their pregnancy. Prenatal checkups can be regarded as an indicator of whether a woman is willing to invest in the pregnancy and is an indicator of other healthy behaviors (Currie and Moretti, 2003).

The preschool reform could have affected health at birth by influencing the decision of women to have fewer children and to have children at older ages. I find that the expansion of preschool facilities increased aged at motherhood and decreased pregnancies, especially among teenagers. This result is in line with a large literature that documents an association between education and fertility choices of women (see Strauss and Thomas, 1995). Women having fewer children could explain why children are born with higher quality. In this chapter I provide evidence that when I exclude the effect on fertility at specific age groups, the observed effect on health at birth remains, suggesting that this channel does not explain the findings.

Interestingly, my findings differ from those in McCrary and Royer (2011), the only other study to date that examines the effect of additional schooling at the beginning of the school trajectory. I find that the effect of girls starting school earlier on health at birth is positive while McCrary and Royer (2011) find that treated and control females give birth to children of similar health. I argue that the improvements in child health may come from increases in prenatal care while McCrary and Royer (2011) do not find any changes in prenatal behavior due to the increased schooling. Several reasons could explain why both studies show different findings. On the one hand, the outcome variables considered are somewhat different. While my study finds intergenerational effects of education on the likelihood of extreme prematurity, McCrary and Royer (2011) only focus on low birth weight and prematurity and find no effects for these margins. On the other hand, the setup in both studies is different. McCrary and Royer (2011) claim that their results may be difficult to generalize to other populations due to specific characteristics of their study. In their setup, the authors focus on mothers that give birth before the age of 23 and that their estimates may disproportionately reflect the experience of women from low socioeconomic backgrounds. In my study I consider slightly older mothers as well as women of broader socioeconomic contexts. Lastly, my analysis pertains to a different country than the one considered in McCrary and Royer (2011). I focus on the case of Uruguay, instead of United States (US), where baseline health at birth outcomes and prenatal behavior are different. The likelihood of low birth weight is higher in my sample (8%) than in the US sample (6%) and first-time mothers in the US sample receive more prenatal care.

3.A Appendix Tables

The tables included in this section supplement the information in the main text. The first table shows the age of females in the sample according to their birth cohort and the year they are observed. The second table shows the availability of preschool places per child by year and region for 5-year-old. The third table shows the number of observations added to the sample to perform the bounding analysis.

Table 3.i: Age of females in the sample by birth cohort and year of observation

Birth cohort	Year									
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
1988	20	21	22	23	24	25	26	27	28	29
1989	19	20	21	22	23	24	25	26	27	28
1990	18	19	20	21	22	23	24	25	26	27
1991	17	18	19	20	21	22	23	24	25	26
1992	16	17	18	19	20	21	22	23	24	25
1993	15	16	17	18	19	20	21	22	23	24
1994	.	15	16	17	18	19	20	21	22	23
1995	.	.	15	16	17	18	19	20	21	22
1996	.	.	.	15	16	17	18	19	20	21

Table 3.ii: Availability of preschool places per child by year and region for 5-year-olds

Region	Year									Increase
	1992	1993	1994	1995	1996	1997	1998	1999	2000	1992-2000
Montevideo	0.50	0.52	0.53	0.51	0.52	0.46	0.51	0.50	0.51	2%
Artigas	0.49	0.50	0.47	0.50	0.49	0.63	0.68	0.62	0.50	3%
Canelones	0.47	0.49	0.53	0.54	0.56	0.56	0.56	0.62	0.60	27%
Cerro Largo	0.67	0.69	0.62	0.67	0.64	0.51	0.73	0.52	0.59	-12%
Colonia	0.81	0.78	0.81	0.78	0.73	0.82	0.79	0.79	0.84	4%
Durazno	0.70	0.62	0.70	0.70	0.88	0.76	0.74	0.55	0.62	-12%
Flores	0.85	0.85	0.91	0.85	0.85	0.90	0.82	0.75	0.75	-12%
Florida	0.89	0.79	0.84	0.89	0.84	0.86	0.79	0.78	0.69	-23%
Lavalleja	0.88	0.88	0.94	0.94	1.06	0.73	0.55	0.73	0.65	-27%
Maldonado	0.38	0.34	0.43	0.48	0.50	0.64	0.47	0.56	0.64	68%
Paysandu	0.43	0.46	0.41	0.47	0.46	0.46	0.58	0.60	0.65	49%
Rio Negro	0.65	0.65	0.73	0.76	0.76	0.64	0.50	0.75	0.69	7%
Rivera	0.59	0.63	0.66	0.66	0.74	0.70	0.79	0.59	0.76	29%
Rocha	0.65	0.70	0.79	0.70	0.77	0.64	0.71	0.67	0.66	1%
Salto	0.44	0.51	0.48	0.48	0.42	0.46	0.45	0.49	0.51	15%
San Jose	0.67	0.57	0.67	0.68	0.76	0.70	0.79	0.57	0.63	-5%
Soriano	0.72	0.72	0.69	0.65	0.80	0.78	0.65	0.64	0.48	-33%
Tacuarembo	0.69	0.76	0.69	0.69	0.80	0.72	0.72	0.74	0.83	-23%
Treinta y Tres	0.70	0.64	0.64	0.79	0.86	0.70	0.68	0.55	0.54	77%

Note: Availability of preschool places per child of age 5 is calculated as the number of groups opened for 5-year olds by region and year multiplied by an average of 25 students per group and divided by the number of children aged 5 in each region in the corresponding year (obtained from the Uruguayan National Institute of Statistics).

Table 3.iii: Number of births added to the sample to perform bounding analysis

Region	Age group		
	16	17	28
1	0	0	0
2	25	30	30
3	44	40	40
4	25	32	32
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	25	11	11
11	28	18	18
12	49	51	51
13	26	21	21
14	0	0	0
15	0	0	0
16	53	57	49
17	0	0	0
18	0	0	0
19	46	48	19

Chapter 4

Long-term Impacts on Education Outcomes of Receiving a Cash Transfer during Early-life[1]

4.1 Introduction

A growing literature documents that prenatal and early-childhood experiences can have long-lasting impacts on later-life economic outcomes, human capital, health and well-being (Almond et al., 2018). In recent years, there has been a large increase of welfare programs that aim to improve conditions in early-life, especially in developing countries. While for policymakers it is particularly interesting to know how effective these policies are, we are still at the beginning stages of learning what type of intervention matters for long-term outcomes.

In this chapter, we evaluate whether being exposed to a poverty-alleviation program during early-life has an impact on long-term educational outcomes. We focus on the Uruguayan *Plan de Atención Nacional a la Emergencia Social* (PANES), a cash transfer program that was implemented between 2005 and 2007 and that targeted the poorest 10 percent of households in the country. The program was introduced after a severe economic crisis that hit Uruguay in 2002. One special feature is that the amount of the transfer represented approximately 45% of the average household income among its target population. Although participation was announced to be conditional on school attendance for all children under age 14 and on regular health checkups for pregnant women and all children, the conditions were never enforced, so the program was unconditional *de facto*.

Evaluating long-term effects of interventions during early childhood has two main challenges. The first challenge is to find a credible identification strategy to eval-

[1]This chapter is based on Bloomfield and Cabrera (2020)

uate the intervention. In this chapter, we exploit the way in which households were assigned to the PANES program. Program assignment was determined on the basis of a baseline predicted poverty score: households whose score was above a certain threshold were eligible to receive the transfer. This eligibility rule generates a discontinuity that we exploit using a regression discontinuity design. We compare educational outcomes of children belonging to households just above and just below the eligibility cutoff. We estimate impacts on three educational outcomes: highest grade attained, delay in educational attainment and dropout.

The second challenge in estimating long-term effects of early-childhood interventions is to find datasets that map early-life environments with later-life outcomes. We make use of a rich dataset that we constructed for this project that links long-term educational outcomes to early life experiences. Our dataset contains educational information (enrollment and grade) for the years 2013-2017 of eligible and ineligible children born between 2003 and 2007.

We separately estimate the impact of the PANES program for cohorts that were exposed at different stages of early childhood. We split our analysis according to the age of the child at the onset of the program. In particular, we focus on (i) children that were between 0 and 2 years of age when the program started (born between January 2003 and March 2005), and on (ii) children that were born during the program period. Therefore, given that the program ran between April 2005 and December 2007, our sample includes children that were exposed to the program between the ages 0 and 5 (those in group (i)) or while in-utero and up to maximum two years and eight months (those in group (ii)).

Separating the analysis into children that were exposed to cash transfers since the in-utero period and children that were exposed later in life (but still in early childhood) enables us to look at differential effects among these subgroups. On the one hand, a growing literature suggests that investments that occur during the prenatal period may potentially be more cost-effective than postnatal interventions (Doyle et al., 2009). In this sense, we should expect that children that received transfers while in-utero and after birth benefit more from the program than those that received transfers only after birth. On the other hand, children born before the start of the program were born during an economic crisis and had more risk: the likelihood of being born with low birth weight was 0.083 in the pre-program period, while 0.075 in the program period. In this sense, we should expect higher effects of the program for these children than for those born in a better economic landscape. Identifying which group of children benefited more from an intervention such as PANES might help focus cash transfer policies on those children that need them most.

Our results show that the PANES program improved educational attainment.

The effect is entirely driven by children exposed to the program during early-childhood. Within this sample, we find that children from eligible households have a higher educational attainment and a lower incidence of delay than ineligible children. In addition, we find that PANES had no impact on educational dropout in the years of observation. Taken together, these findings suggest that the effect of PANES on educational attainment works through retainment and not through dropout.

Following Heckman's model of dynamic complementarity, one would expect that children that received transfers since the in-utero period should have stronger effects on education than those that received them only after birth. However, our findings show the opposite, the effect of the PANES program on education is driven by children exposed to the program during early childhood (and after birth). Because this group was born in a worse economic environment, our interpretation is that the transfers have a stronger effect on education on children that are born with more risk. We further explore this issue by estimating heterogeneous effects by low birth weight status among children that were born in the pre-program period. We find that the effects of PANES on educational attainment are stronger among children that were born with low birth weight.

We find no effects of PANES on long-term educational results of children that were in their mothers' womb during the program period. Given previous evidence showing that the PANES program improved health at birth as measured by birth weight (Amarante et al., 2016)[2] and the importance of health at birth for later educational outcomes,[3] this finding seems surprising. However, when we estimate the effect of PANES on health at birth, we find no significant effects. Our findings differ from those in Amarante et al. (2016) because we use a different identification strategy and a different dataset.[4]

Cash transfer programs have been very popular in developing countries, particularly in Latin America. These welfare programs were established with the aim of

[2]The authors find that the PANES program led to a drop in the incidence of low birth weight that ranges between 19 and 25 percent and that fertility was not affected by program participation. The result could be considered a "first stage" effect for our long-term educational outcomes, although the program may affect long-term outcomes also through other mechanisms (Almond et al., 2018)

[3]Birth weight has emerged as the main focus of health policy, both in the United States and elsewhere, and has been used to evaluate the effectiveness of social policy (Almond et al., 2005). Research has shown that birth weight can affect neonatal outcomes and long-run health outcomes (Black et al., 2007; Oreopoulos et al., 2008), and even birth weight of the next generation (Royer, 2009; Black et al., 2007). Birth weight can also affect non-health outcomes such as schooling, wages, IQ and test scores (Behrman and Rosenzweig, 2004; Royer, 2009; Oreopoulos et al., 2008; Rosenzweig and Zhang, 2013; Black et al., 2007; Torche and Echevarría, 2011).

[4]Amarante et al. (2016) use a localized difference in differences strategy while we use a regression discontinuity design. When Amarante et al. (2016) use a regression discontinuity design, they do not find significant impacts of PANES on low birth weight. When we perform a localized difference in differences strategy, we do not find robust results showing that the program improved health at birth. We discuss this issue further in Section 4.5.3.

alleviating household financial restrictions. Some variations of these programs impose conditionalities on school attendance to promote human capital accumulation and break the intergenerational transmission of poverty. Unconditional cash transfers have been shown to increase household consumption (Haushofer and Shapiro, 2016) and educational attainment in the short-run (Baird et al., 2013) but their effectiveness in improving the outcomes associated with conditions is inferior relative to conditional cash transfers (Baird et al., 2011). While there is sufficient evidence of the impact of cash transfers in the short-run (Fiszbein and Schady, 2009), the evidence on long-run effects is sparse (Millán et al., 2019). In general, most studies find that cash transfers have positive effects on schooling while impacts on employment and earnings are mixed (see Millán et al. (2019) for a review).[5]

In Uruguay, the PANES program has been evaluated on a range of short-term outcomes such as school attendance, labor supply, political support and birth weight. Overall, studies find that the program had no impact on child labor or school attendance of children aged 14 to 17 (Amarante et al., 2013), decreased formal labor supply (Amarante et al., 2011), increased political support for the current government relative to the previous government (Manacorda et al., 2011) and improved health at birth outcomes (Amarante et al., 2016).

Our chapter contributes to the literature in two ways. First, we contribute to a growing body of work on the medium to long-term effects of (unconditional) cash transfer programs in developing countries. We measure educational outcomes 8 to 12 years after exposure, a longer period than that in most other studies. Second, beyond the cash transfer literature, we contribute to the literature that relates resources in-utero to educational outcomes later in life. While most other studies have focused on long-term effects of negative shocks such as famines, disease and radiation (see Almond et al. (2018) for a recent review), we focus on a policy that implies a positive treatment.

The remainder of this chapter is structured as follows. Section 4.2 describes the PANES program, Sections 4.3 and 4.4 describe the data and empirical framework respectively. Section 4.5 reports results of the effect of the PANES program on educational outcomes and low birth weight. Finally, Section 4.6 provides a discussion of the findings.

[5]Some studies find no or little long-term impacts of cash transfers on education. Two examples are Haushofer and Shapiro (2018) and Blattman et al. (2020). The former evaluates an unconditional cash transfer in Kenya three years after the beginning of the program and the latter evaluates the effectiveness of cash grants in Uganda 9 years after the implementation of the program.

4.2 The PANES Program

The *Plan de Atención Nacional a la Emergencia Social* (PANES) was a temporary social assistance program that ran between April 2005 and December 2007, in Uruguay, a middle-income country in Latin America.[6] The program targeted the poorest households in the country. The PANES was designed as an emergency plan to alleviate material hardship from a severe economic crisis that hit Uruguay in 2002 and was among the flagship policies of the center-left government that took office in March 2005. The Ministry for Social Development (*Ministerio de Desarrollo Social*) was created to be in charge of the implementation of the program.

Program eligibility was based on families' scores on a poverty index. All applicant households were visited by personnel of the Ministry of Social Development and completed a detailed baseline survey which allowed program officials to compute the score. The score depended on many household socioeconomic characteristics and was based on a probit model of the likelihood of being below a per capita income level using a highly saturated function of household variables (Amarante et al., 2005). The estimation of the underlying model was performed using the 2003 and 2004 National Household Survey (*Encuesta Continua de Hogares*) and the resulting coefficient estimates were used to predict a score value for each applicant using PANES baseline survey data. Appendix Tables 4.i and 4.ii provide further information on the variables used to predict the poverty score.[7] The variables considered, the weights attached to the observed covariates and the eligibility thresholds were allowed to vary slightly across different geographic regions. Applicants were not aware of the variables that entered into the score, nor the weights attached to them, or the eligibility criterion, easing concerns about manipulation of the score.

Rather than using actual reported income, the score was estimated using a wide range of socioeconomic variables. The reason for this is that the program's target population often worked in the informal sector making it difficult to verify self-reported income. By using indirect measures of income the possibility of strategic misreporting was minimized.

Around 188,671 households (with around 700,000 individuals) sent applications. After the interviewing process, households were ordered according to their level of deprivation based on their predicted poverty score. Those households whose score was above a predetermined level were assigned to the program. Around 54% of

[6] In 2003, Uruguay had a population of around 3.3 million people and per capita GDP was about 8000 USD. The country offers free public education from elementary school to university. There are 14 years of mandatory schooling: 2 in elementary school, 6 in primary school and 5 in secondary school. While primary education is universal, secondary school completion rates pose a big challenge for the Uruguayan government.

[7] One of the variables used to predict the poverty score was the household's value in a wealth index. The variables included in the latter measure are listed in Appendix Table 4.ii.

applicant households became beneficiaries, representing nearly 10% of households in the country. Independently of their characteristics, eligible households received a monthly cash transfer that originally amounted to $1360 Uruguayan pesos (US$102 adjusted by PPP). This amount was adjusted for inflation on a quarterly basis. The transfer corresponded to approximately 45 percent of the average household income among the poorest 10 percent of households in Uruguay.[8] [9]

The condition to keep receiving the payment was that household income (of all sources) remained below a specific level per capita. In practice, only verifiable sources of income were taken into account. Successive checks were carried out by the social security administration to enforce this condition and, because of this, some households stopped receiving the transfer before the end of the program.[10] There were no other formal conditionalities (such as health checks for children and pregnant women or school attendance for children) until mid 2007, and even then, conditionalities were never enforced.

The program included several components. The main element of the program was the monthly cash transfer (*ingreso ciudadano*, "citizen income"). Midway through the program, an electronic food card (*tarjeta alimentaria*) was introduced and households with children or pregnant women were entitled to receive it on top of the cash transfer. The food card operated through an electronic debit card and its value represented between 22% and 59% of the value of the income transfer depending on household size and demographic structure.

On an annual basis, the program's cost was 0.41% of GDP. The program ended in December 2007 and the target population, eligibility rules and assistance levels changed when a new system of family allowances and a health care reform (*Plan de Equidad*) was launched in January 2008. Households did not need to reapply for the new program. The eligibility to the *Plan de Equidad* was based on a new score that was estimated for all original PANES applicant households using the same baseline characteristics registered in 2005 but with a new formula. The threshold for program eligibility changed with respect to PANES: it became less restrictive

[8]This number was calculated using the Uruguayan Continuous Household Survey of 2004. If we use the wave of 2005, we obtain very similar results. Income is substantially lower outside Montevideo, the capital city of Uruguay, which explains why 70% of applicants live outside the capital city. The fixed amount of $1360 Uruguayan pesos represent slightly more than 50% of monthly average household income among the poorest 10 percent households that do not live in Montevideo and slightly less than 40% of monthly average household income among the poorest 10 percent households that live in Montevideo. With respect to the whole income distribution of the country, the transfer represents a 9% of the monthly household average income.

[9]Our calculations are line with Amarante and Vigorito (2010) and Amarante et al. (2011) who state that the monthly amount of the transfer corresponded to half (50%) of the pre-program household self-reported income. In Amarante et al. (2016), the authors state that the amount of the transfer represented a quarter of self-reported income (25%).

[10]Households that became non-eligible before the end of the program are still considered within the treatment group.

and expanded the beneficiaries' base.[11] The government informed households about the ending of the PANES and the start of the new program via mail and eligible households received a written formal communication.

4.3 Data

We use a rich dataset that links administrative records from three governmental sources. All sources contain information at the individual level and we use de-identified identity numbers for matching these three sources. In this section we describe the data sources used and the descriptive statistics.

4.3.1 Data Sources

Data from the Ministry of Social Development

Our primary source is the administrative records of the initial baseline survey visit for both successful and unsuccessful female applicants in PANES. The Ministry of Social Development (*Ministerio de Desarrollo Social*, MIDES) shared with us the responses to the comprehensive questionnaire applied by MIDES agents during the visits. Some households submitted more than one application to the program but we keep information only from the first visit to ease concerns about strategic behaviors to gain eligibility. The key variables that we use from this source are the household's exact predicted poverty score and an indicator for approval in PANES. We also use information on the household's sociodemographic characteristics, housing conditions and durable asset ownership.

Birth data

We combine information from PANES administrative records with all registered live births in Uruguay coming from birth certificates (*Certificado de Nacido Vivo*) in the period 2003-2007. The latter are registered by the Statistical Office of the Ministry of Health (*Ministerio de Salud Pública*). Birth certificates have unique identification numbers for mothers and we used these to match them with females in PANES applicant households. The identity numbers of children, however, were not available in birth certificates. To obtain this information, we used additional

[11] Members from eligible and ineligible households in PANES became eligible for the new program. In 79% of applicant households to PANES, at least one household member became eligible in *Plan de Equidad*. Further in the chapter we show that we do not find significant differences in the probability that at least one household member received the *Plan de Equidad* when considering households that are close to the PANES eligibility threshold (See Table 4.3). It is important to note, however, that in this chapter we estimate the marginal effect of receiving the PANES program on top of receiving future cash transfers from *Plan de Equidad*.

records of MIDES that contain identification numbers of mothers and children that receive any social program. We matched the latter dataset with PANES records using the mother's identification number and the date of birth. For multiple births of the same gender, it was not possible to disentangle which was the identification number that corresponded to each child. Because this information was key to link birth data with education data, we had to drop observations from multiple births (1% of the sample).[12] The vital statistics natality data has information of health at birth, the reproductive history of the mother, parental characteristics and prenatal health care utilization.

Education data

Finally, we use children's identification numbers to obtain information of enrollment by year and grade from administrative data registered by the Statistical Office of the National Administration of Public Education (*Administración Nacional de Educación Pública*). We have information for the years 2013 to 2017,[13] corresponding to 8 to 12 years after the beginning of the PANES program. For each year, we know the grade in which the child was enrolled but not whether the grade was completed in that particular year. With this data, we constructed three outcome variables for our analysis: highest grade attained, delay and dropout in education. Highest grade attained corresponds to the grade attained by the child in 2017, the last year for which we have information, and it ranges from 1 to 10 being 1 preschool and 10 the last year of middle school. If the observation of the child is missing in 2017, we take the highest grade attained by the child in the period we observe her.[14] Delay is measured with an indicator that takes value 1 if the child's highest grade attained in 2017 is lower than the one determined by her year and month of birth and a regular track.[15] Appendix Table 4.iii shows the corresponding grade that a child should have attained in 2017 according to its year and month of birth. Dropout is an indicator that takes the value 1 if the child was not enrolled in education for two or more years during the period of observation.

[12]Within the program period, multiple births are equally likely for PANES recipients as for controls, easing concerns of selection on an outcome. Infants born in multiple births have, on average, lower birth weights than those born in single order births, so our results may be sensible to the inclusion of twins, triplets and higher order births.

[13]The data that is used in this project was collected in 2018. We would like to update the evaluation period but we have not been able to get access to more recent data yet.

[14]We acknowledge that we do not measure completed education and that highest grade attained is a truncated variable. We have performed our estimations using an alternative outcome variable that measures the likelihood of enrolling in sixth grade with no delay which excludes the possibility of truncation for younger students. The results are qualitatively equivalent to the ones we show in our main tables (see Table 4.vii).

[15]In Uruguay, the requirement to enter the public education system is to have the age corresponding to the level before April 30 of the school year. That causes most children (2/3) to reach the age following the level during the school year and that 1/3 of the children do it the other year.

The three outcomes we consider capture different elements of students' educational career. Highest grade attained shows overall educational attainment of the child. Delay adds to the latter by considering also information of the year and month of birth of the child. There are two possible explanations to why a child may be enrolled at a lower grade than the one we would expect her to be based on her age and a regular track: (i) the child repeated a grade, or (ii) the child did not enroll in school during some years.[16] We explore the possibility of explanation (ii) using a variable that indicates whether the child dropped out from school for two or more years in the period we observe her.

4.3.2 Descriptive Statistics

Overall, we have information of 49,062 mothers and 59,128 children. Almost half of the children in our sample (49%) were born during the program period. Table 4.1 presents descriptive statistics of our outcome variables and selected covariates for children born in the pre-program period (January 2003-March 2005) and children born in the program period (April 2005-December 2007). There is a difference in educational outcomes measured 8 to 12 years after exposure to the program between eligible and ineligible groups. For example, taking into account highest grade attained, children born in non-eligible households in the pre-program period attained 8 years of education while eligible children born in the same period attained 7.6. Children born during the program period attain lower levels of education than children born in the pre-program period because they are younger at the time we observe them. Non-eligible children born after the beginning of the program attain on average 5.8 years of education while eligible children attain 5.7.

There is also a difference in the incidence of low birth weight between eligible and ineligible households. In the pre-program period, 8.7% of eligible children were born with low birth weight while among ineligible children the incidence was 7.9%. During the program period, the gradient in low birth weight is less pronounced (7.7% and 7.4% for eligible and ineligible households respectively).

[16] A third explanation could be that the parents delayed the enrollment of the child at the first grade of education. We are not able to capture this as a separate outcome because we do not observe the full educational trajectory of the child and therefore we do not know in which year they entered school. In out setting, having parents that enroll children at a higher cohort than the one they should enter could be problematic in terms of our outcome measures because if these kids repeat a grade we would still consider them as non-delayed. Even though age cutoffs to enter preschool are not strictly enforced in Uruguay, the children that enroll early are minority, and it is more common to see children enrolled late instead.

Table 4.1: Descriptive statistics of outcome variables and selected covariates

	Eligible households		Non eligible households		Difference	
	N	Mean	N	Mean	Coefficient	s.e.
Panel A: Born in pre-program period						
Child's highest grade attained	22157	7.758	6751	7.975	−0.217***	(0.014)
Delay	22157	0.602	6,751	0.482	0.120***	(0.007)
Dropout	22157	0.025	6751	0.031	−0.006**	(0.002)
Mother's number of previous pregnancies	22157	3.381	6751	2.362	1.019***	(0.030)
Mother's age at birth	21778	25.432	6667	24.433	0.998***	(0.093)
Child's birth weight (BW) in grams	22000	3175.765	6696	3199.972	−24.206***	(7.217)
Child has low birth weight (=1 if BW<2500 grams)	22000	0.087	6696	0.079	0.008**	(0.004)
Gestational week of birth occurrence	21512	38.637	6581	38.647	−0.010	(0.025)
Child was born premature (=1 if gestational weeks<37)	21512	0.082	6,581	0.081	0.001	(0.004)
Child's APGAR score 1 minute	21923	8.535	6709	8.550	−0.015	(0.014)
Child's APGAR score 5 minute	21929	9.642	6708	9.649	−0.007	(0.011)
Mother's number of prenatal controls	21940	6.560	6678	7.560	−0.961***	(0.046)
Panel B: Born during program period						
Child's highest grade attained	22221	5.717	7999	5.823	−0.106***	(0.014)
Delay	22221	0.382	7999	0.284	0.097***	(0.006)
Dropout	22221	0.020	7999	0.032	−0.012***	(0.002)
Mother's number of previous pregnancies	22221	3.396	7999	2.422	0.974***	(0.028)
Mother's age at birth	21772	25.124	7873	24.675	0.449***	(0.088)
Child's birth weight (BW) in grams	22012	3214.934	7925	3225.874	−10.940	(6.857)
Child has low birth weight (=1 if BW<2500 grams)	22012	0.077	7925	0.074	0.003	(0.003)
Gestational week of birth occurrence	21353	38.648	7654	38.664	−0.015	(0.023)
Child was born premature (=1 if gestational weeks<37)	21353	0.079	7654	0.078	0.001	(0.004)
Child's APGAR score 1 minute	22005	8.538	7925	8.516	0.021*	(0.013)
Child's APGAR score 5 minute	22004	9.651	7925	9.641	0.010	(0.009)
Mother's number of prenatal controls	21887	6.777	7869	7.728	−0.951***	(0.043)

Note: Standard errors (s.e.) are reported in parentheses. N corresponds to number of observations. * p<.1, ** p<.05, *** p<.01.

4.4 Empirical Framework

In this section we explain the main identification strategy used to estimate the impacts of the PANES program on long-term educational results. We also show first-stage estimates of the effect of eligibility in the PANES transfer on actual treatment and discuss the validity of our main estimates.

4.4.1 Identifying Long-run Impacts of the PANES Program

To examine the impact of the PANES program on educational attainment 8 to 12 years after exposure to the program, we use a regression discontinuity design. We exploit the fact that program assignment was determined by a predicted poverty score. Families that ranked above a certain threshold were eligible to receive the cash transfer while those below the threshold were not. This rule creates a discontinuity in the probability of receiving the transfer. Given that eligibility enforcement is high but not perfect, we estimate program effects using a fuzzy regression discontinuity design.

We compare outcomes of children that were born in households that were just above and just below the cutoff. The equation that we estimate is the following:

$$Y_{imt} = \alpha_0 + \alpha_1 T_m + f(N_m) + \alpha_2 X_{imt} + e_{imt} \tag{4.1}$$

where Y is the schooling outcome of interest of child i conceived by mother m and born in year t, T_m is a binary indicator variable that takes the value 1 if the mother m received the benefit or 0 otherwise, N_m denotes mother m's predicted poverty score (normalized relative to the eligibility threshold such that households with positive N_m are eligible for treatment), f is a function of the running variable that is continuous at the threshold (N_m=0) and that may have different slopes at each side of the cutoff. All regressions control for month times year of birth fixed effects, and month times year of baseline visit fixed effects. X_{imt} include the latter fixed effects and may also include other controls as we mention in the following paragraph. e_{imt} is a random error term. We instrument the PANES treatment variable T_m, with an indicator for the mother's program eligibility, E_m. α_1 is the parameter of interest.

As in fully randomized experiments, it is not necessary to include covariates in regression discontinuity designs. However, it is often the case that studies include them to reduce variability in the estimation (Lee and Lemieux, 2010). In our estimations we control for covariates, X_{imt}, at the level of the child, of the mother and of the household.[17] Controls are included as indicator variables and we use a

[17]We control for covariates that are not used to predict the eligibility score (see Tables 4.i and

separate category for missing observations in each control.

We estimate Equation 4.1 for the entire sample and on two subsamples: children exposed to the cash transfer while they were in-utero (i.e. those born between May 2005 and December 2007) and children exposed to the cash transfer after birth (i.e. those born between January 2003 and April 2005). We report results based on the bandwidth and polynomial selected following the approach of Calonico et al. (2014).[18] This approach consists of a local polynomial nonparametric estimator with data-driven bandwidth selector and biased-correction techniques. We refer to this approach as "CCT". In most cases, the optimal bandwidth ranges between 0.05 and 0.1 (meaning, respectively, differences of 5 to 10 percentage points in the predicted poverty score).

One concern is that pregnancy might be endogenous to gaining program eligibility. Having one more child would increase the probability of treatment since the score was estimated using the *per capita* income level of the household. This could bias the estimates of program impact if women who change their pregnancy patterns give birth to children with different characteristics, for example, with a different probability of low birth weight. Given that the initial application period was concentrated in a relatively short period of time (75% of applications took place in the first nine months of the program), it seems unlikely that in such period fertility patterns may have been influenced. A related issue is the possibility of any fertility responses to the program in order to retain eligibility. To ease concerns about later fertility choices, we use the predicted income score at the initial application as an instrument for program receipt, instead of the score at each reassessment of eligibility status (where circumstances in the household, including child birth, may have changed).

4.4.2 First-stage Effects of the PANES Program

Figure 4.1 shows a clear jump in the fraction of individuals that actually received the PANES transfers.[19] While 96% of poor households located to the right of the cutoff received the cash transfer, 13% of ineligible households managed to enter the program.

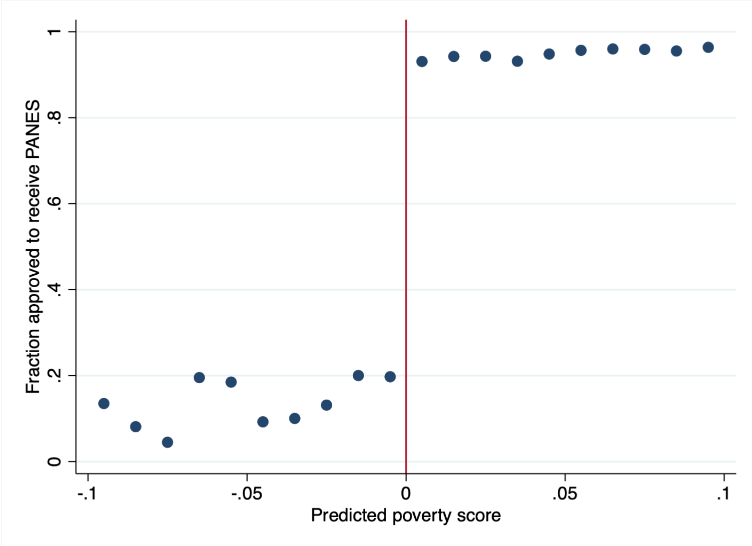
Table 4.2 presents first-stage estimates of the effect of eligibility in the PANES transfer on actual treatment. We report results using three different ranges around the eligibility threshold (Columns (1)-(8)). We also report results for the bandwidth

4.ii) with the exception of those that are unbalanced at baseline. In particular, we control for: gender of the child, educational level of the mother, indicators for whether the household's block has sewage, trash collection and for the number of rooms in the household.

[18]Calonico et al. (2014) incorporates the latests advances in regression discontinuity methods and refines the estimator proposed by Imbens and Kalyanaraman (2012)

[19]Note that the normalized predicted poverty score ranges from -0.19 to 0.95 in our sample.

Figure 4.1: Receipt of PANES



Note: The vertical line corresponds to the eligibility cutoff, above which households are eligible to the program and below which they are not eligible to the program. There are 10 bins at each side of the cutoff and the range is -0.1, 0.1. Each dot represents the fraction of households that received the PANES transfers in that bin.

defined according to Calonico et al. (2014) (Columns (13)-(14)). In Panel A we report estimates for the whole sample, in Panel B we report estimates for children exposed during early childhood and in Panel C we report estimates for those exposed while in-utero. The estimated increase in the fraction of treated households at the threshold is large (between 0.70 and 0.76) and does not change much between specifications.[20] The first-stage estimates become larger when using observations that are further away from the cutoff.

[20]We obtain very similar results when using a second order polynomial function (see Table 4.iv in the Appendix).

Table 4.2: First stage estimates of the effect of the eligibility on the PANES cash transfer

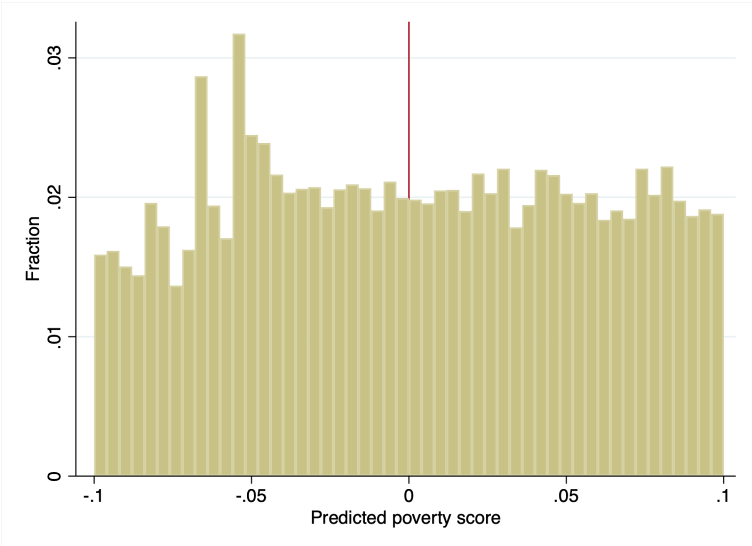
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All observations								
Coefficient	0.747***	0.755***	0.741***	0.747***	0.701***	0.717***	0.702***	0.716***
s.e.	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)	(0.013)	(0.012)
Observations	25622	25622	19863	19863	13262	13262	9358	9592
Range	0.1	0.1	0.075	0.075	0.05	0.05	0.036	0.037
Panel B: Exposed during early-childhood								
Coefficient	0.742***	0.757***	0.736***	0.750***	0.697***	0.720***	0.693***	0.718***
s.e.	(0.010)	(0.010)	(0.012)	(0.011)	(0.014)	(0.014)	(0.019)	(0.018)
Observations	12198	12198	9435	9435	6277	6277	5119	5156
Range	0.1	0.1	0.075	0.075	0.05	0.05	0.042	0.042
Panel C: Exposed while in-utero								
Coefficient	0.751***	0.754***	0.744***	0.747***	0.703***	0.715***	0.707***	0.710***
s.e.	(0.009)	(0.009)	(0.011)	(0.011)	(0.014)	(0.013)	(0.019)	(0.018)
Observations	13424	13424	10428	10428	6985	6985	5066	5053
Range	0.1	0.1	0.075	0.075	0.05	0.05	0.037	0.037
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Each cell corresponds to a different regression. In Columns (1)-(6) we estimate Equation 4.1 using as outcome variable an indicator that takes the value of one if the household received the PANES transfer. We report results for three different fixed ranges around the eligibility threshold and a first order polynomial. We also report the estimates obtained when using the bandwidth and polynomial defined according to Calonico et al. (2014) (Columns (7)-(8)). All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Estimations in even columns we include the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. All controls are included as indicator variables and include a category for missing observations. Standard errors (s.e.) are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

4.4.3 Testing the Identifying Assumptions

The regression discontinuity design assumes that assignment to either side of the threshold is as good as random. To check whether there is bunching just above or just below the threshold, we plot a density graph of the running variable (predicted poverty score) for the whole sample (Figure 4.2) and for each of the two subsamples (Figure 4.vii and Figure 4.viii in the Appendix). A visual inspection of the density graphs suggests that bunching does not occur. More formally, we test bunching by conducting a McCrary's density test (McCrary, 2008) using observations near the threshold.[21] The log difference in height is 0.022 (s.e. 0.047) in the full sample, 0.019 (s.e. 0.060) in the sample of children exposed during early-childhood and 0.041 (s.e. 0.066) in the sample of children exposed while in-utero.

Figure 4.2: Density



Note: The figure shows the distribution in the range of -0.1 and 0.1 of the running variable. Each bar represents the fraction of households in specific values of the predicted poverty score. The vertical line corresponds to the eligibility cutoff.

To check whether covariates are balanced at baseline, we run Equation 4.1 using a wide range of baseline household, mother and child characteristics.[22] Table 4.3 reports results from estimating the effect of the PANES program on the different covariates at baseline. Most coefficients are small and not significantly different from zero which is in line with assignment around the threshold being as good as random.

[21]We use observations that have a value of the running variable in the range -0.1 and 0.1.

[22]We use pre-program data for those covariates that are not measured at baseline: birth weight, low birth weight, apgar 1 minute, apgar 5 minutes, age of the mother at birth, number of prenatal controls, gestational weeks and number of previous pregnancies.

Moreover, a joint significance test gives a p-value of 0.159.[23] Most covariates have a strong correlation with highest grade attained, yet they are balanced between treated and controls (see Column (5)).[24] Boys have a lower educational attainment than girls, being born with low birth weight has a negative correlation with highest grade attained and children whose mothers have completed primary education attain higher grades than those with lower educated mothers.

We include estimates of the effect of the PANES transfer on different covariates for children exposed to the program during early-childhood and for children exposed to the program while in-utero separately in Tables 4.v and 4.vi in the Appendix. Balancing in the sample of children exposed while in-utero is similar to that in the full sample, with only three coefficients showing-up significant. This is consistent with the identification assumption that assignment around the threshold is as good as random. In the case of the sample of children exposed to PANES during early-childhood, coefficients are significant in a few more cases but the sign of these coefficients go in the opposite direction of the correlation of the covariate with the outcome highest grade attained. In any case, we control for all pre-treatment covariates

[23] The estimation is performed using pre-program data and considers the optimal bandwidth obtained in Table 4.2.

[24] We checked these correlations for the other educational outcomes and the conclusion is the same.

Table 4.3: Estimates of the effect of the PANES transfer on different covariates using baseline data and correlation of covariates with main outcome

	Non-eligible mean	Coefficient	s.e.	N	Correlation with outcome
	(1)	(2)	(3)	(4)	(6)
Child's indicators					
Child is a boy	0.510	0.008	(0.014)	19863	−0.191∗∗∗(0.011)
Birth weight	3199	18.840	(22.140)	9362	0.000∗∗∗(0.000)
Low birth weight	0.080	-0.007	(0.012)	9362	−0.180∗∗∗(0.033)
Apgar 1 minute	8.575	-0.037	(0.043)	9368	0.026∗∗∗(0.009)
Apgar 5 minutes	9.675	0.002	(0.031)	9366	0.044∗∗∗(0.013)
Age in months in Dec 2007	29.930	-0.307	(0.490)	19863	0.021∗∗∗(0.001)
Mother's indicators					
Age	24.627	0.005	(0.278)	9290	−0.003∗∗∗(0.001)
Complete primary education	0.925	0.014*	(0.008)	19826	0.341∗∗∗(0.020)
Complete secondary education	0.031	0.001	(0.005)	19826	0.130∗∗∗(0.033)
Number of prenatal controls	7.463	0.218	(0.135)	9349	0.026∗∗∗(0.003)
Gestational weeks	38.640	0.073	(0.077)	9169	0.026∗∗∗(0.005)
Number of previous pregnancies	2.564	-0.219***	(0.076)	9435	−0.046∗∗∗(0.005)
Household's indicators					
Hot water	0.294	0.019	(0.012)	19859	0.081∗∗∗(0.013)
Heater	0.192	0.001	(0.011)	19845	0.090∗∗∗(0.014)
Kitchen	0.684	0.012	(0.014)	19862	0.113∗∗∗(0.012)
Heating	0.007	-0.003	(0.002)	19833	−0.023(0.071)
Concrete floor	0.555	-0.012	(0.014)	19653	−0.050∗∗∗(0.011)
Mud wall	0.920	0.010	(0.008)	19551	0.109∗∗∗(0.019)
Block has electricity	0.978	0.000	(0.005)	19858	0.067∗(0.035)
Block has piped water	0.940	0.010	(0.007)	19851	0.060∗∗∗(0.023)
Block has sewage	0.409	0.050***	(0.014)	19793	0.058∗∗∗(0.011)
Block has trash collection	0.900	0.017*	(0.009)	19835	0.092∗∗∗(0.018)
Block has paved streets	0.666	0.006	(0.014)	19797	0.032∗∗∗(0.012)
Block has sidewalk	0.701	0.009	(0.013)	19808	0.077∗∗∗(0.012)
House	0.879	-0.015	(0.010)	19533	0.068∗∗∗(0.016)
Microwave	0.045	0.002	(0.005)	19863	0.111∗∗∗(0.029)
Refrigerator	0.662	0.018	(0.014)	19848	0.065∗∗∗(0.011)
Freezer	0.092	0.008	(0.008)	19824	0.035∗(0.020)
Washing machine	0.186	0.000	(0.011)	19863	0.021(0.015)
Dishwasher	0.002	0.001	(0.001)	19849	0.065(0.128)
TV	0.791	0.005	(0.012)	19859	0.104∗∗∗(0.013)
VCR	0.040	0.008	(0.005)	19857	−0.018(0.030)
Cable TV	0.134	0.010	(0.009)	19863	0.122∗∗∗(0.018)
Computer	0.010	0.001	(0.003)	19855	0.143∗∗(0.063)
Car	0.031	0.003	(0.005)	19863	0.080∗∗(0.035)
Home owned	0.498	-0.006	(0.014)	19831	−0.041∗∗∗(0.011)
Number of rooms	2.407	0.059	(0.055)	19861	0.016∗∗∗(0.003)
Number of bedrooms	1.721	0.047**	(0.024)	19861	0.022∗∗∗(0.007)
Receipt of *Plan de Equidad*	0.804	-0.006	(0.011)	19392	0.065∗∗∗(0.014)

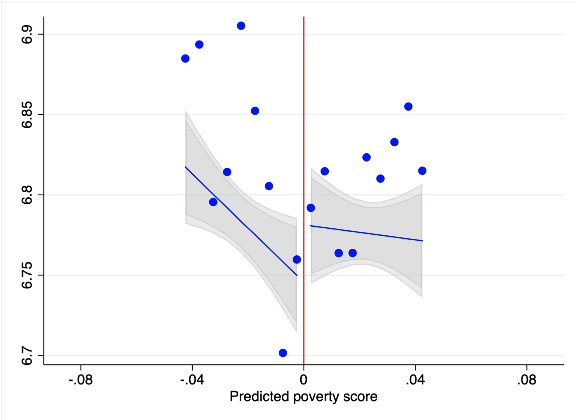
Note: In Column (2) we report estimates of Equation 4.1 using different covariates at baseline as outcome variables. We use pre-program data for those covariates that are not measured at baseline: birth weight, low birth weight, apgar 1 minute, apgar 5 minutes, age of the mother at birth, number of prenatal controls, gestational weeks and number of previous pregnancies. Estimates are obtained using a bandwidth of 0.075 around the threshold and a first order polynomial. In Column (5) we report the correlation of each covariate with the outcome highest grade attained. We obtain these correlations by regressing highest grade attained on each covariate and conditioning on month times year of birth fixed effects, and month times year of baseline visit fixed effects. Standard errors (s.e.) are reported in parentheses. N corresponds to number of observations. * p<.1, ** p<.05, *** p<.01.

4.5 Empirical Results

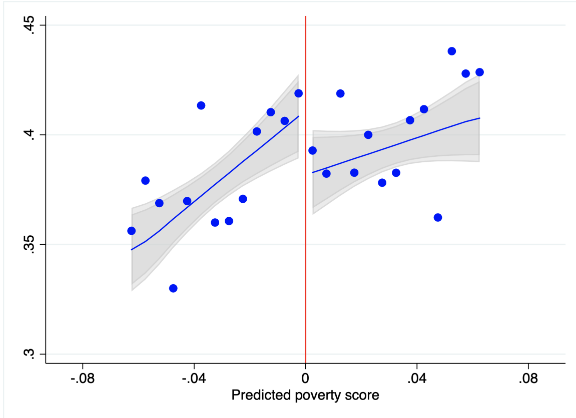
4.5.1 Main Results

In Figures 4.3-4.5 we plot highest grade attained (in Panel A), delay in educational attainment (in Panel B) and dropout (in Panel C) for different values of the predicted poverty score. The lines are average values for each outcome from linear regressions fitted to the sample of children whose predicted poverty score was within the optimal bandwidth defined by Calonico et al. (2014) around the eligibility cutoff (0). The light grey and dark grey shaded areas correspond to 90% and 95% confidence bands respectively. Figure 4.3 shows educational outcomes for the pooled sample of children in our study. Children's highest grade attained decreases as the poverty score increases, the incidence of delay increases as the poverty score increases and dropout increases as the poverty score increases. In the three panels there is a clear jump at the threshold that suggests that barely eligible children have better educational outcomes than barely ineligible children. However, the only jump that seems to be significant is the one in Panel B, at the 10% level.

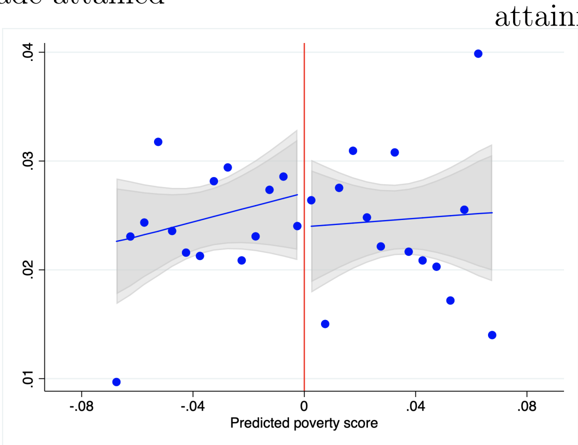
Figure 4.3: Educational outcomes of pooled sample of children



Panel A: Highest grade attained



Panel B: Delay in educational attainment



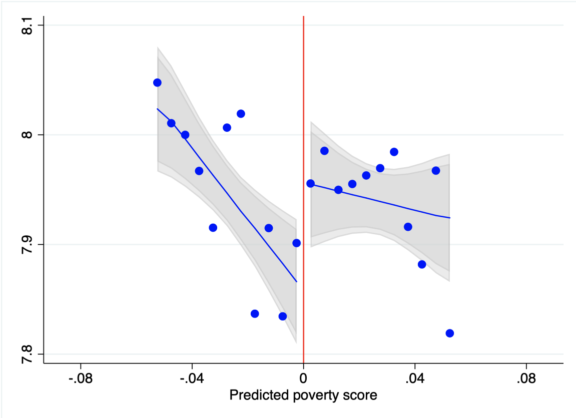
Panel C: Dropout

Note: The vertical line corresponds to the eligibility cutoff, above which households are eligible to the program and below which they are not eligible to the program. Each dot represents the average outcome in a bin. The two solid lines represent the best fit from a linear regression from each side of the cut-off using observations within the optimal bandwidth defined by Calonico et al. (2014). The light grey and dark grey shaded areas correspond to 90% and 95% confidence bands respectively.

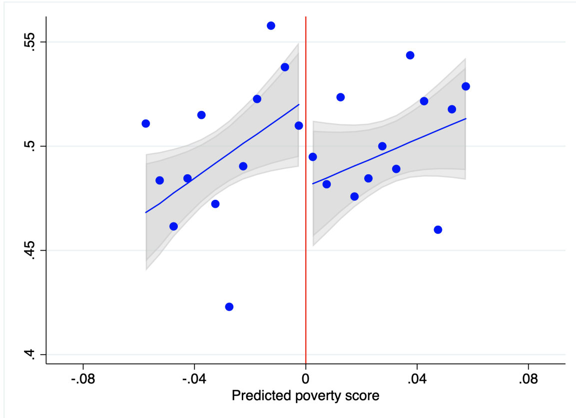
Figure 4.4 shows outcomes of children born in the pre-program period, hence, those exposed to the program during early childhood. Barely eligible children have a higher educational attainment and a lower incidence of delay than barely ineligible children (see the jump at the threshold in Panels A and B). The jump in highest grade attained seems to be significant at the 10% level. A visual inspection of the plot in Panel C suggests that the likelihood of dropout is higher for eligible children than for ineligible children but this difference is not significant.[25]

[25] The observed dropout rate is consistent with the fact that primary education is nearly universal in Uruguay. 98% of children that enroll in the first year of primary school complete primary education.

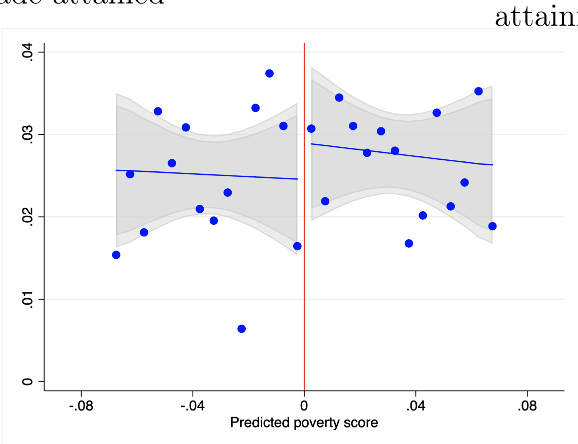
Figure 4.4: Educational outcomes of children exposed to the program during early-childhood



Panel A: Highest grade attained



Panel B: Delay in educational attainment

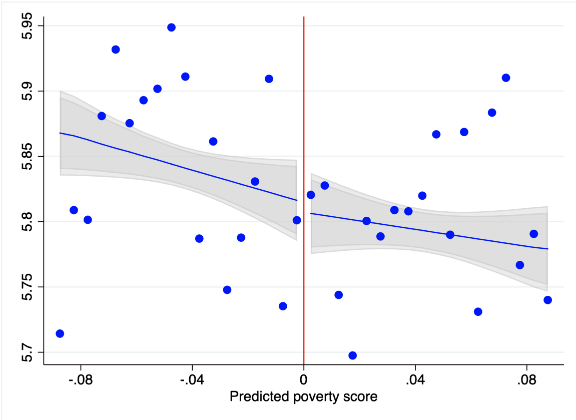


Panel C: Dropout

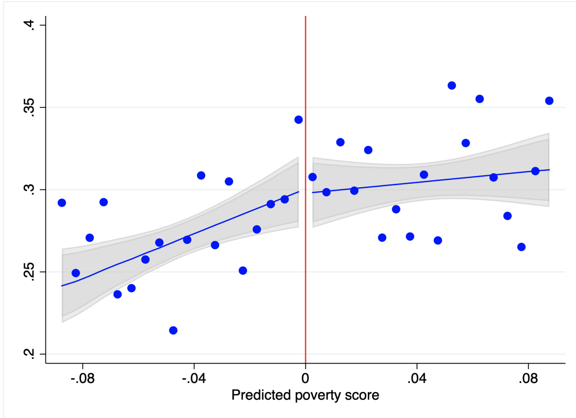
Note: The vertical line corresponds to the eligibility cutoff, above which households are eligible to the program and below which they are not eligible to the program. Each dot represents the average outcome in a bin. The two solid lines represent the best fit from a linear regression from each side of the cut-off using observations within the optimal bandwidth defined by Calonico et al. (2014). The light grey and dark grey shaded areas correspond to 90% and 95% confidence bands respectively.

Figure 4.5 shows educational outcomes of children exposed to the PANES program while in-utero. In Panel C there is a jump at the threshold suggesting that eligible children have lower dropout. However, this jump is non-significant. In Panels A and B we observe practically no jump at the threshold.

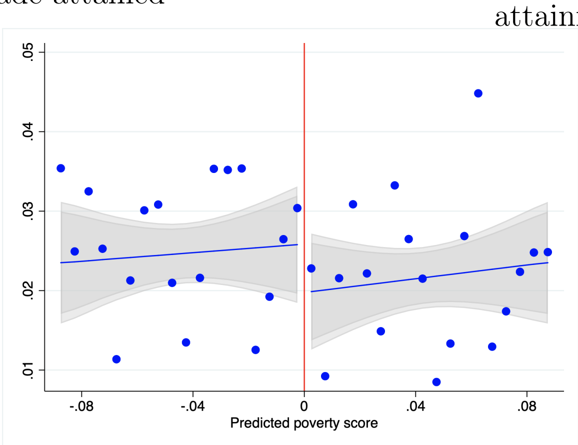
Figure 4.5: Educational outcomes of children exposed to the program while in-utero



Panel A: Highest grade attained



Panel B: Delay in educational attainment



Panel C: Dropout

Note: The vertical line corresponds to the eligibility cutoff, above which households are eligible to the program and below which they are not eligible to the program. Each dot represents the average outcome in a bin. The two solid lines represent the best fit from a linear regression from each side of the cut-off using observations within the optimal bandwidth defined by Calonico et al. (2014). The light grey and dark grey shaded areas correspond to 90% and 95% confidence bands respectively.

In Table 4.4 we report estimates of the effect of receiving the PANES transfer during early childhood and while in-utero on educational attainment 8 to 12 years later. In Panel A we report estimates for the whole sample, in Panel B we report estimates for children exposed during early childhood and in Panel C we report estimates for those exposed while in-utero. For each outcome, we use two specifications one with controls and one without controls.[26] By and large, coefficients go in the expected direction: the effects on highest grade attained should be positive,

[26]Note that the number of observations in each regression changes according to the bandwidth. For the same sample, the number of observations changes whether we use or not use controls. These changes not always go in the same direction. In Tables 4.viii and 4.ix in the Appendix we report results using specific bandwidths with a fixed number of observations.

the effects of delay in educational attainment should be negative and the effects on dropout should be negative. For the entire sample (Panel A), there is a negative and significant effect on the probability of being delayed (the p-value is 0.07 and 0.08 in Columns (3) and (4) respectively). When splitting the sample, we find that eligible children that were exposed to the program during early-childhood (Panel B) have a higher educational attainment (the p-value in Columns (1) and (2) is 0.08). In addition, we find that the effect on educational attainment is due to a lower incidence of delay in education (the p-value is 0.06 and 0.08 in Columns (3) and (4) respectively). We find no significant effects on dropout. The latter comes at no surprise given that we are considering children that are mainly in primary school and dropout is more likely to occur in secondary. Educational attainment of children exposed to the program while in-utero (Panel C) is not significantly different between eligible and ineligible households in any of the outcomes considered.

Table 4.4: Effect of receiving the PANES transfer on educational attainment

	Highest grade attained		Delay		Dropout	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All observations						
Coefficient	0.043	0.047	-0.040*	-0.038*	-0.001	-0.002
s.e.	(0.044)	(0.040)	(0.022)	(0.022)	(0.007)	(0.007)
Observations	11613	13586	16597	16744	18224	18297
CCT bandwidth	0.045	0.051	0.062	0.063	0.068	0.068
Panel B: Exposed during early-childhood						
Coefficient	0.125*	0.112*	-0.069*	-0.064*	0.008	0.007
s.e.	(0.071)	(0.065)	(0.036)	(0.037)	(0.011)	(0.010)
Observations	6553	7376	7608	7264	8350	9111
CCT bandwidth	0.052	0.058	0.060	0.057	0.066	0.072
Panel C: Exposed while in-utero						
Coefficient	-0.001	-0.002	-0.015	-0.012	-0.009	-0.009
s.e.	(0.034)	(0.034)	(0.024)	(0.023)	(0.008)	(0.008)
Observations	11917	11809	12065	12453	11973	11667
CCT bandwidth	0.087	0.086	0.088	0.092	0.088	0.084
Controls	No	Yes	No	Yes	No	Yes

Note: Each cell corresponds to a different regression. In Panel A we use the sample of children whose family received the PANES transfer during early-childhood. In Panel B we use the sample of children whose family received the PANES transfer while the child was in-utero. In Panel C we use all observations. We estimate Equation 4.1 using different outcome variables. In Columns (1)-(2) we report results using as outcome variable highest grade attained in education. In Columns (3)-(4) we report results using as outcome variable an indicator for delay in educational attainment that takes value 1 if the child is enrolled at a lower grade than the one determined by her year and month of birth and a regular track. In Columns (5)-(6) we report results using as outcome variable an indicator for whether the child dropped out from education, where dropout is measured as being two or more years not enrolled in any program. We report the estimates obtained when using the CCT bandwidth and polynomial defined according to Calonico et al. (2014). CCT bandwidths are reported below each coefficient and CCT polynomial is 1 in all cases. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. In even columns we present results of estimating Equation 4.1 using the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. All controls are included as indicator variables and include a category for missing observations. Standard errors (s.e.) are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

4.5.2 Heterogeneous Effects by Low Birth Weight

Children who were first exposed to PANES during early childhood were born at the time of a severe economic crisis. In this subsection we explore whether transfers have a stronger effect on education on children that are born with more risk of weighting less than 2500 grams. Table 4.5 shows heterogeneous effects of receiving the PANES transfer on education by low birth weight. Estimations only consider observations of children born in the pre-program period, and hence, that were exposed to the program during early childhood. The first two columns correspond to children that were born with low birth weight (=1 if $<$2500 grams) and the last two columns correspond to children that were born with normal birth weight. Barely eligible children that are born with a weight less than 2500 grams have a higher educational attainment and a lower incidence of delay than barely ineligible children born with the same condition. There is also an effect on dropout but goes in the opposite direction than expected. In particular, the likelihood of dropout is higher among barely eligible children born with low birth weight than among barely ineligible children born with low birth weight. The latter effect is significant at the 10 percent level. Overall, the findings from this analysis suggest that the program had stronger effects on children who were born when the economic context was more unfavorable and with more risk of low birth weight.

4.5.3 Exploring Short-run Impacts of PANES on Low Birth Weight

We found no evidence supporting that the PANES program improved educational attainment for those exposed while in-utero. At a first glance, these results are surprising given the large literature on the effects of low birth weight on educational attainment (Figlio et al., 2014) and previous evidence showing that PANES improved health at birth. In this subsection we explore whether low birth weight is a potential mechanism behind long-term educational outcomes.

We use a regression discontinuity approach and compare health at birth outcomes between eligible and ineligible children that were born during the program period. A visual inspection of the incidence of low birth weight at both sides of the PANES eligibility cutoff (Figure 4.6) suggests that the program had no impact on health at birth. Table 4.6 shows results from estimating Equation 4.1 using low birth weight as the outcome variable. We find that the relevant coefficients are negative but are small in magnitude and not significant. Our findings are in line with Buser et al. (2017) which finds no effect on weight and height of gaining a cash transfer in Ecuador. Our conclusion is that low birth weight cannot be considered a first stage effect for our long-term impacts on education of children exposed to the program

Table 4.5: Heterogeneous effects of receiving the PANES transfer on education by low birth weight

	Born with low birth weight		Born with normal birth weight	
	(1)	(2)	(3)	(4)
Panel A: Highest grade attained				
Coefficient	0.500**	0.529***	0.078	0.070
s.e.	(0.199)	(0.198)	(0.067)	(0.058)
Observations	783	761	6936	8434
CCT bandwidth	0.075	0.072	0.060	0.074
Panel B: Delay				
Coefficient	-0.241**	-0.255***	-0.050	-0.040
s.e.	(0.099)	(0.099)	(0.035)	(0.035)
Observations	865	848	7923	7657
CCT bandwidth	0.084	0.082	0.068	0.067
Panel C: Dropout				
Coefficient	0.065*	0.069*	0.004	0.003
s.e.	(0.039)	(0.041)	(0.011)	(0.010)
Observations	844	789	8591	9259
CCT bandwidth	0.081	0.076	0.075	0.082
Controls	No	Yes	No	Yes

Note: Each cell corresponds to a different regression. Estimations consider the sample of children whose family received the PANES transfer during early-childhood. We estimate Equation 4.1 using different outcome variables and for two different subsamples. Each panel corresponds to a different outcome. In Columns (1)-(2) we report results for children with low birth weight. In Columns (3)-(4) we report results for children without low birth weight. We report the estimates obtained when using the CCT bandwidth and polynomial defined according to Calonico et al. (2014). CCT bandwidths are reported below each coefficient and CCT polynomial is 1 in all cases. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. In even columns we present results of estimating Equation 4.1 using the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. All controls are included as indicator variables and include a category for missing observations. Standard errors (s.e.) are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

while in-utero.

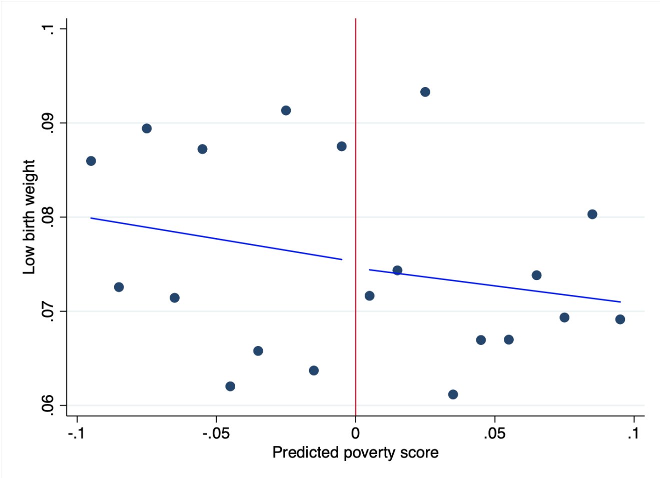
Following Amarante et al. (2016), we also report results on health at birth using a localized difference in differences estimator (See Appendix B for details of this identification strategy).[27] Table 4.7 reports results from estimating Equation 4.2 (see Appendix B) using low birth weight as outcome variable.[28] From our estimations, we cannot reject the null hypothesis of no effect of the PANES transfer on the incidence of low birth weight. The estimated coefficients are negative and mostly non-significant. Standard errors increase as we get closer to the cut-off.[29] These

[27] This method was first formalized by Grembi et al. (2016) but others have executed similar empirical strategies in prior literature. Grembi et al. (2016) propose and verify a set of diagnostic tests for this design. They refer to this method as "difference in discontinuity design". Identification rests on the difference between two cross-sectional estimators instead of within unit variation in treatment assignment.

[28] We include an equivalent set of control variables as those used in Amarante et al. (2016).

[29] We do not use a CCT bandwidth for these estimations given that the equation we estimate does not correspond to a traditional regression discontinuity design.

Figure 4.6: Low birth weight around the PANES cutoff



Note: The vertical line corresponds to the eligibility cutoff, above which households are eligible to the program and below which they are not eligible to the program. There are 10 bins at each side of the cutoff and the range is -0.1, 0.1. Each dot represents the average low birth weight in that bin. The two solid lines represent the best fit from a linear regression from each side of the cut-off.

Table 4.6: 2SLS estimates of the effect of the PANES transfer on low birth weight (<2.500 kg) children born during program period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. No controls	-0.001	-0.006	0.000	-0.008	-0.009	-0.008	-0.003
	(0.012)	(0.019)	(0.014)	(0.027)	(0.018)	(0.027)	(0.013)
Observations	13283	13283	10309	6905	6905	6905	14243
Bandwidth	0.1	0.1	0.075	0.075	0.05	0.05	0.114
Order of polynomial	1	2	1	2	1	2	1
2. Controls	-0.001	-0.005	0.001	-0.006	-0.009	-0.006	-0.003
	(0.012)	(0.019)	(0.014)	(0.027)	(0.018)	(0.027)	(0.013)
Observations	13283	13283	10309	6905	6905	6905	13384
Bandwidth	0.1	0.1	0.075	0.075	0.05	0.05	0.101
Order of polynomial	1	2	1	2	1	2	1

Note: Each cell corresponds to a different regression. Sample includes children that were born during the program period. In Columns (1)-(6) we estimate Equation 4.1 for three different bandwidths around the eligibility threshold and two different orders of polynomial. In Column (7) we report the estimates obtained when using the bandwidth and polynomial defined according to Calonico et al. (2014). All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Row 1 presents regressions with no additional controls while row 2 reports results with the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. Controls are included as indicator variables and include a category for missing observations. Standard errors are reported in parentheses and number of observations are reported below each coefficient. * $p<.1$, ** $p<.05$, *** $p<.01$.

results differ with the findings from Amarante et al. (2016)[30] and we attribute this to the fact that the sample that we use is different. Specifically, our database does not include multiple births[31] which on average are those that are born with lower weights.[32] In any case, as we discuss further in Appendix B, we cannot validate all of the assumptions of the localized difference in differences in the setting of this chapter, hence, our regression discontinuity estimates are our preferred specification.

Table 4.7: 2SLS estimates of the effect of the PANES transfer on low birth weight (<2.500 kg) using a difference in discontinuity design

	(1)	(2)	(3)	(4)	(5)	(6)
1. No controls	-0.011*	-0.011*	-0.013	-0.013	-0.014	-0.014
	(0.006)	(0.006)	(0.008)	(0.008)	(0.010)	(0.010)
Observations	56856	56856	25384	25384	19670	19670
2. Controls	-0.013**	-0.013**	-0.013	-0.013	-0.013	-0.013
	(0.006)	(0.006)	(0.008)	(0.008)	(0.010)	(0.010)
Observations	56856	56856	25384	25384	19670	19670
Range	All	All	0.1	0.1	0.075	0.075
Order of polynomial	1	2	1	2	1	2

Note: Each cell corresponds to a different regression. Sample contains pooled pre-program and program period data, corresponding to children born between the years 2003 and 2007. In Columns (1)-(6) we estimate Equation 4.2 for three different ranges around the eligibility threshold and three different orders of polynomial. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Row 1 presents regressions with no additional controls while row 2 reports results with the following additional controls: gender of the child, number of previous pregnancies of the mother, an indicator for whether the mother completed primary school, indicators for geographic department of the household at baseline, for whether the household has centralized hot water, heater, kitchen, microwave, refrigerator, freezer, washing machine, dishwasher, TV, VCR, cable TV, computer, car, whether the block has electricity, piped water, sewage, trash collection, paved streets, sidewalk, whether the home is a house, is owned, and indicators for material of the floor and walls. Controls are included as indicator variables and include a category for missing observations. Standard errors are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

[30] Note, however, that Amarante et al. (2016) do not find an effect of PANES on low birth weight when using a regression discontinuity strategy.

[31] We acknowledge that not having data for multiple births is somewhat unfortunate, in future research we would like to replicate our analysis using this additional data.

[32] Amarante et al. (2016) find that exposure to PANES reduces the incidence of birth weights below 3000 grams and that effects grow at lower birth weights.

4.6 Discussion

There is relatively little evidence on long-term effects of early-childhood programs. In this chapter we explore whether expanding economic resources during early-life in the form of a cash transfer improves later outcomes. In particular, we explore the effect of being exposed to the Uruguayan PANES in the prenatal period and during early childhood on educational outcomes 8 to 12 years later. We use a rich dataset that matches administrative data from three sources and enables us to distinguish effects for children that were exposed since they were in their mother's womb and children that were exposed to the program in the first years of life.

Our results show that children from eligible households that started receiving the program after they were born, have a higher educational attainment and a lower likelihood of educational delay. We find no impacts on dropout which suggests that the increase in educational attainment probably works through a reduction in the likelihood of repeating a grade. This is in line with the findings of other studies showing that cash transfers have a positive impact on school progression (Millán et al., 2019).

Our effects are concentrated among children that were born before the program period, and not on those exposed in the in-utero period. We find an increase on educational attainment of 0.1 years of education and a decrease in the likelihood of delay of 6.9 percentage points around the eligibility cutoff. These results correspond to local average treatment effects around the cutoff point. Considering that the amount of the transfer represented almost half of the average household income among its population, the magnitude of the effects of PANES on education is rather small.

One potential explanation to why we find results for the subsample of relatively older children and not on the relatively younger children is that the former sample were born and started receiving the program when the Uruguayan context was more unfavorable and poverty rates were higher. Note that total income in PANES applicant families doubled between the pre-program and program period. Our interpretation is that, more than arguing against Heckman's theory of dynamic complementarities, our findings suggest that the program has an impact on children born in families that are close to the eligibility cutoff when children are born in a worse economic situation, and with more risk of low birth weight.

4.A Appendix Tables and Figures

This section includes several tables and figures to supplement the information in the main text. Tables and figures show: (i) the variables that enter the poverty score, (ii) the corresponding grade of children in the sample according to their birth date, (iii) first stage estimates using a second order polynomial, (iv) the bunching and balancing properties of each subsample, (v) the effect of the PANES program on the likelihood of enrolling in sixth grade without delay, and (vi) 2SLS estimates of the effect of the PANES program for fixed bandwidths.

Table 4.i: Variables included in the poverty score

	Urban areas		Rural areas
	Capital city	Other regions	
Public employees in the household	✓	✓	
Retirees in the household	✓	✓	✓
Pensioners in the household	✓	✓	
Logarithm of the number of household members	✓	✓	✓
Presence of children aged 0-5	✓	✓	
Presence of adolescents aged 12-17	✓	✓	
Presence of children aged 0-4			✓
Presence of children aged 5-10			✓
Presence of adolescents aged 11-17			✓
Wealth index (See Table 4.ii)	✓	✓	✓
Average years of education of adults	✓	✓	
Household's head completed primary education			✓
Residential overcrowding	✓	✓	✓
Toilet facilities: no toilet	✓		
Toilet facilities: flush toilet	✓		
Toilet facilities: pit latrine	✓		
Toilet facilities: other	✓		
Toilet facilities: no toilet		✓	
Toilet facilities: flush toilet or pit latrine		✓	
Toilet facilities: other		✓	
Toilet facilities: no cistern			✓
Masonry			✓
Concrete floor			✓
Dirt floor			✓
House is owned	✓		
House is leased	✓		
House is occupied	✓		
Household type: head only			✓
Household type: head and spouse			✓
Household type: head and children			✓
Household type: head, spouse and children only			✓
Household type: head, spouse, children and other relatives			✓
Household type: head, spouse, children and other non-relatives			✓
At least one of the household's member has mutual insurance			
Household's head has mutual insurance			✓
Year	✓	✓	
Constant	✓	✓	

Note: Own elaboration based on Amarante et al. (2005). The model used to predict the poverty score was estimated using the Continuous Household Survey of 2003 and 2004.

Table 4.ii: Variables used to construct the wealth index

	Urban areas	Rural areas
Ownership of water heater	✓	✓
Ownership of boiler	✓	
Ownership of fridge	✓	✓
Ownership of color television	✓	✓
Access to cable television	✓	
Ownership of videocassette recorder	✓	✓
Ownership of washing machine	✓	✓
Ownership of dishwasher	✓	
Ownership of microwave	✓	
Ownership of laptop computer	✓	
Ownership of car	✓	✓
Ownership of telephone	✓	✓

Note: Own elaboration based on Amarante et al. (2005).

Table 4.iii: Corresponding grade in 2017 according to child's year and month of birth

		Month of birth											
		January	February	March	April	May	June	July	August	September	October	November	December
	2003	10	10	10	10	9	9	9	9	9	9	9	9
	2004	9	9	9	9	8	8	8	8	8	8	8	8
Year of birth	2005	8	8	8	8	7	7	7	7	7	7	7	7
	2006	7	7	7	7	6	6	6	6	6	6	6	6
	2007	6	6	6	6	5	5	5	5	5	5	5	5

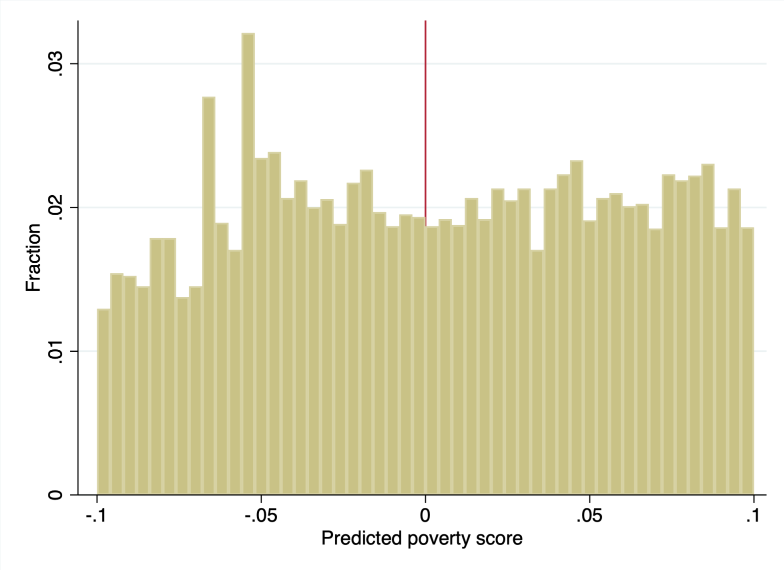
Note: Table shows corresponding grade that a children should attain according to its year and month of birth in Uruguay. Grade 1 corresponds to the last year of preschool education and grade 10 corresponds to the third year of secondary school. The requirement to enter the Uruguayan public education system is to have the age corresponding to the level before April 30 of the school year. That causes most children (2/3) to reach the age following the level during the school year and that 1/3 of the children do it the other year

Table 4.iv: First stage estimates of the effect of the eligibility on the PANES cash transfer using a second order polynomial function

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All observations						
Coefficient	0.747***	0.755***	0.741***	0.747***	0.702***	0.718***
s.e.	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)
Observations	25622	25622	19863	19863	13262	13262
Range	0.1	0.1	0.075	0.075	0.05	0.05
Panel B: Exposed during early-childhood						
Coefficient	0.741***	0.756***	0.736***	0.749***	0.698***	0.720***
s.e.	(0.010)	(0.010)	(0.012)	(0.011)	(0.014)	(0.014)
Observations	12198	12198	9435	9435	6277	6277
Range	0.1	0.1	0.075	0.075	0.05	0.05
Panel C: Exposed while in-utero						
Coefficient	0.751***	0.754***	0.744***	0.747***	0.704***	0.716***
s.e.	(0.009)	(0.009)	(0.011)	(0.011)	(0.014)	(0.013)
Observations	13424	13424	10428	10428	6985	6985
Range	0.1	0.1	0.075	0.075	0.05	0.05
Controls	No	Yes	No	Yes	No	Yes

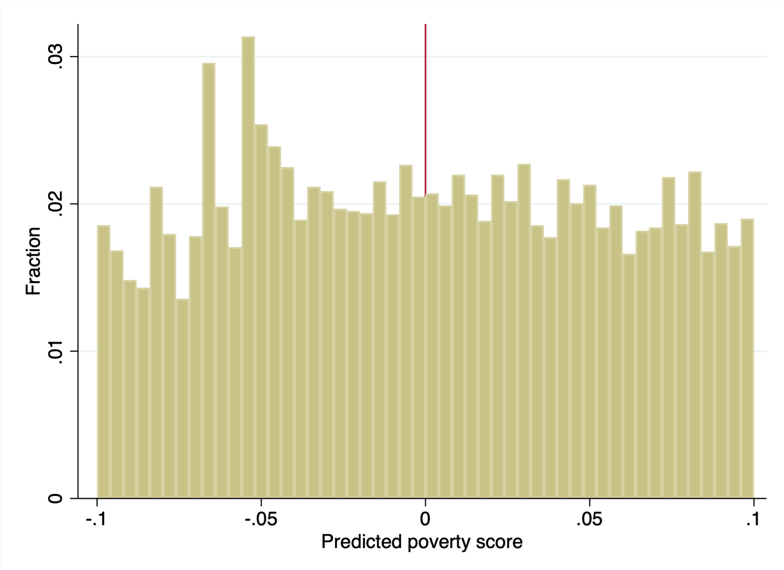
Note: Each cell corresponds to a different regression. In Columns (1)-(6) we estimate Equation 4.1 using as outcome variable an indicator that takes the value of one if the household received the PANES transfer. We report results for three different fixed ranges around the eligibility threshold and a second order polynomial. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Estimations in even columns we include the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. All controls are included as indicator variables and include a category for missing observations. Standard errors (s.e.) are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

Figure 4.vii: Density for subsample of children exposed during early-childhood



Note: The figure shows the distribution in the range of -0.1 and 0.1 of the running variable for the subsample of children exposed during early childhood. Each bar represents the fraction of households in specific values of the predicted poverty score. The vertical line corresponds to the eligibility cutoff.

Figure 4.viii: Density for subsample of children exposed while in-utero



Note: The figure shows the distribution in the range of -0.1 and 0.1 of the running variable for the subsample of children exposed while in-utero. Each bar represents the fraction of households in specific values of the predicted poverty score. The vertical line corresponds to the eligibility cutoff.

Table 4.v: Estimates of the effect of the PANES transfer on different covariates using baseline data for subsample of children exposed during early-childhood

	Non-eligible mean	Coefficient	s.e.	N
	(1)	(2)	(3)	(4)
Child's indicators				
Child is a boy	0.511	0.010	(0.021)	9435
Age in months in Dec 2007	33.368	-1.014	(0.730)	9435
Mother's indicators				
Age	24.627	0.005	(0.278)	9290
Complete primary education	0.932	0.008	(0.011)	9422
Complete secondary education	0.036	0.003	(0.008)	9422
Household's indicators				
Hot water	0.313	0.043**	(0.019)	9434
Heater	0.196	-0.007	(0.017)	9428
Kitchen	0.696	0.024	(0.020)	9435
Heating	0.007	-0.003	(0.003)	9421
Concrete floor	0.928	0.018	(0.012)	9287
Mud wall	0.537	-0.010	(0.021)	9330
Block has electricity	0.978	0.004	(0.007)	9433
Block has piped water	0.937	0.009	(0.010)	9431
Block has sewage	0.417	0.044**	(0.021)	9402
Block has trash collection	0.894	0.015	(0.013)	9422
Block has paved streets	0.664	0.037*	(0.020)	9403
Block has sidewalk	0.703	0.023	(0.019)	9411
House	0.884	-0.006	(0.014)	9273
Microwave	0.046	0.007	(0.008)	9435
Refrigerator	0.679	0.012	(0.020)	9426
Freezer	0.093	0.015	(0.012)	9412
Washing machine	0.197	0.029*	(0.016)	9435
Dishwasher	0.002	0.004**	(0.002)	9429
TV	0.797	0.019	(0.017)	9434
VCR	0.040	0.009	(0.008)	9434
Cable TV	0.134	0.021	(0.013)	9435
Computer	0.010	0.004	(0.004)	9433
Car	0.032	0.004	(0.007)	9435
Home owned	0.500	-0.007	(0.021)	9421
Number of rooms	2.391	0.159**	(0.073)	9434
Number of bedrooms	1.715	0.093***	(0.035)	9434
Receipt of *Plan de Equidad*	0.871	-0.023	(0.015)	9287

Note: In Column (2) we report estimates of Equation 4.1 using different covariates at baseline as outcome variables for the subsample of children exposed during early-childhood. Estimates are obtained using a bandwidth of 0.075 around the threshold and a first order polynomial. Standard errors (s.e.) are reported in parentheses. N corresponds to number of observations. * p<.1, ** p<.05, *** p<.01.

Table 4.vi: Estimates of the effect of the PANES transfer on different covariates using baseline data for subsample of children exposed while in-utero

	Non-eligible mean	Coefficient	s.e.	N
	(1)	(2)	(3)	(4)
Child's indicators				
Child is a boy	0.509	0.007	(0.020)	10428
Age in months in Dec 2007	26.885	0.553	(0.638)	10428
Mother's indicators				
Age	24.765	0.012	(0.258)	10242
Complete primary education	0.919	0.018	(0.011)	10404
Complete secondary education	0.027	-0.001	(0.006)	10404
Household's indicators				
Hot water	0.278	-0.001	(0.017)	10425
Heater	0.188	0.008	(0.015)	10417
Kitchen	0.674	0.001	(0.019)	10427
Heating	0.007	-0.002	(0.003)	10412
Concrete floor	0.913	0.003	(0.012)	10264
Mud wall	0.571	-0.013	(0.020)	10323
Block has electricity	0.978	-0.003	(0.006)	10425
Block has piped water	0.943	0.011	(0.009)	10420
Block has sewage	0.401	0.055***	(0.019)	10391
Block has trash collection	0.905	0.019	(0.012)	10413
Block has paved streets	0.668	-0.021	(0.019)	10394
Block has sidewalk	0.700	-0.003	(0.018)	10397
House	0.875	-0.023*	(0.014)	10260
Microwave	0.045	-0.002	(0.007)	10428
Refrigerator	0.646	0.023	(0.019)	10422
Freezer	0.092	0.002	(0.011)	10412
Washing machine	0.177	-0.026*	(0.014)	10428
Dishwasher	0.002	-0.001	(0.002)	10420
TV	0.785	-0.006	(0.017)	10425
VCR	0.040	0.008	(0.007)	10423
Cable TV	0.134	0.000	(0.012)	10428
Computer	0.009	-0.002	(0.003)	10422
Car	0.029	0.003	(0.006)	10428
Home owned	0.496	-0.006	(0.020)	10410
Number of rooms	2.420	-0.038	(0.080)	10427
Number of bedrooms	1.730	0.004	(0.032)	10427
Receipt of *Plan de Equidad*	0.744	0.010	(0.176)	10105

Note: In Column (2) we report estimates of Equation 4.1 using different covariates at baseline as outcome variables for the subsample of children exposed while in-utero. Estimates are obtained using a bandwidth of 0.075 around the threshold and a first order polynomial. Standard errors (s.e.) are reported in parentheses. N corresponds to number of observations. * p<.1, ** p<.05, *** p<.01.

Table 4.vii: Effect of receiving the PANES transfer on the likelihood of enrolling in sixth grade of primary school without delay

	Likelihood of enrolling in sixth grade with no delay	
	(1)	(2)
Panel A: All observations		
Coefficient	0.058*	0.057**
s.e.	(0.033)	(0.029)
Observations	7552	9720
CCT bandwidth	0.043	0.054
Panel B: Exposed during early childhood		
Coefficient	0.083**	0.072**
s.e.	(0.039)	(0.035)
Observations	5518	6640
CCT bandwidth	0.045	0.053
Panel C: Exposed while in-utero		
Coefficient	0.011	0.030
s.e.	(0.056)	(0.054)
Observations	2454	2635
CCT bandwidth	0.047	0.050
Controls	No	Yes

Note: Each cell corresponds to a different regression. In Panel A we use the sample of children whose family received the PANES transfer during early-childhood. In Panel B we use the sample of children whose family received the PANES transfer while the child was in-utero. In Panel C we use all observations. We estimate Equation 4.1 using the likelihood of enrolling in sixth grade of primary school with no delay as outcome variable. We report the estimates obtained when using the CCT bandwidth and polynomial defined according to Calonico et al. (2014). CCT bandwidths are reported below each coefficient and CCT polynomial is 1 in all cases. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. In Column (2) we present results of estimating Equation 4.1 using the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. All controls are included as indicator variables and include a category for missing observations. Standard errors (s.e.) are reported in parentheses and number of observations are reported below each coefficient. * p<.1, ** p<.05, *** p<.01.

Table 4.viii: 2SLS estimates of the effect of the PANES transfer during early-childhood on educational outcomes for fixed bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Highest grade attained						
1. No controls	0.043	0.045	0.117**	0.117**	0.118*	0.117*
	(0.043)	(0.043)	(0.050)	(0.050)	(0.064)	(0.064)
Observations	12198	12198	9435	9435	6277	6277
2. Controls	0.053	0.128**	0.128**	0.137**	0.135**	0.024**
	(0.044)	(0.051)	(0.051)	(0.065)	(0.065)	(0.010)
Observations	12198	12198	9435	9435	6277	6277
Panel B: Delay in educational attainment						
1. No controls	-0.034	-0.035	-0.067**	-0.067**	-0.056	-0.056
	(0.024)	(0.024)	(0.028)	(0.028)	(0.036)	(0.036)
Observations	12198	12198	9435	9435	6277	6277
2. Controls	-0.040	-0.073***	-0.073***	-0.066*	-0.065*	-0.006*
	(0.024)	(0.028)	(0.028)	(0.036)	(0.036)	(0.003)
Observations	12198	12198	9435	9435	6277	6277
Panel C: Dropout						
1. No controls	0.014*	0.014*	0.003	0.003	0.007	0.007
	(0.008)	(0.008)	(0.009)	(0.009)	(0.012)	(0.012)
Observations	12198	12198	9435	9435	6277	6277
2. Controls	0.014*	0.003	0.003	0.008	0.008	-0.005
	(0.008)	(0.009)	(0.009)	(0.012)	(0.012)	(0.010)
Observations	12198	12198	9435	9435	6277	6277
Range	0.1	0.1	0.075	0.075	0.05	0.05
Order of polynomial	1	2	1	2	1	2

Note: Each cell corresponds to a different regression. Sample includes children born before the program period. Each Panel corresponds to a different outcome. In Columns (1)-(6) we estimate Equation 4.1 using Ordinary Least Squares. We use three different ranges around the eligibility threshold and two different orders of polynomial. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Row 1 presents regressions with no additional controls while row 2 reports results with the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. Controls are included as indicator variables and include a category for missing observations. Standard errors are reported in parentheses and number of observations are reported below each coefficient. * $p<.1$, ** $p<.05$, *** $p<.01$.

Table 4.ix: 2SLS estimates of the effect of the PANES transfer while in-utero on educational outcomes for fixed bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Highest grade attained						
1. No controls	-0.039	-0.039	-0.001	-0.001	-0.006	-0.007
	(0.029)	(0.029)	(0.033)	(0.033)	(0.042)	(0.042)
Observations	13424	13424	10428	10428	6985	6985
2. Controls	-0.039	0.007	0.007	-0.012	-0.014	0.007
	(0.029)	(0.034)	(0.034)	(0.043)	(0.043)	(0.009)
Observations	13424	13424	10428	10428	6985	6985
Panel B: Delay in educational attainment						
1. No controls	0.021	0.021	-0.010	-0.010	-0.012	-0.011
	(0.020)	(0.020)	(0.023)	(0.023)	(0.030)	(0.030)
Observations	13424	13424	10428	10428	6985	6985
2. Controls	0.020	-0.016	-0.016	-0.008	-0.007	-0.009***
	(0.020)	(0.024)	(0.024)	(0.030)	(0.030)	(0.003)
Observations	13424	13424	10428	10428	6985	6985
Panel C: Dropout						
1. No controls	-0.008	-0.008	-0.010	-0.010	-0.008	-0.008
	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)
Observations	13424	13424	10428	10428	6985	6985
2. Controls	-0.008	-0.010	-0.010	-0.009	-0.009	0.012
	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)	(0.015)
Observations	13424	13424	10428	10428	6985	6985
Range	0.1	0.1	0.075	0.075	0.05	0.05
Order of polynomial	1	2	1	2	1	2

Note: Each cell corresponds to a different regression. Sample includes births that occurred in the program period. Each Panel corresponds to a different outcome. In Columns (1)-(6) we estimate Equation 4.1 using Ordinary Least Squares. We use three different ranges around the eligibility threshold and two different orders of polynomial. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Row 1 presents regressions with no additional controls while row 2 reports results with the following additional controls: gender of the child, an indicator for whether the mother completed primary school, number of rooms in the household and indicators for whether the household's block has sewage and trash collection. Controls are included as indicator variables and include a category for missing observations. Standard errors are reported in parentheses and number of observations are reported below each coefficient. * $p<.1$, ** $p<.05$, *** $p<.01$.

4.B Results on Low Birth Weight Using a Difference in Discontinuity Design

Following Amarante et al. (2016), in this chapter we report results of estimating the effect of the PANES program on low birth weight using a localized difference in differences estimator. Amarante et al. (2016) implement this methodology to add more observations to the estimation and improve precision. Since we can observe health at birth outcomes of children born during the pre-program period (before May 2005) and during the program period (between May 2005 and December 2007) in our data, we can take the difference of the local average treatment effect between pre and post treatment discontinuities. More specifically, we focus on changes in outcomes among eligible versus ineligible mothers/children across the pre-program and program period within a close neighborhood of the eligibility threshold. The estimator is then:

$$(E[Y|T_m = 1, D_{imt} = 1] - E[Y|T_m = 0, D_{imt} = 1]) \\ - (E[Y|T_m = 1, D_{imt} = 0] - E[Y|T_m = 0, D_{imt} = 0])$$

where D_{imt} is an indicator for births that took place during the program period and $E[Y|T_m = 1, D_{imt} = 1]$ is the average outcome for children born during the program period in a treated household, $E[Y|T_m = 0, D_{imt} = 1]$ is the average outcome for children born during the program period in a control household, $E[Y|T_m = 1, D_{imt} = 0]$ is the average outcome for children born before the program period in a treated household, and $E[Y|T_m = 0, D_{imt} = 0]$ is the average outcome for children born before the program period in a control household.

To implement this, we estimate the following regression with instrumental variables:

$$Y_{imt} = \beta_0 + \beta_1 D_{imt} + \beta_2 T_m + \beta_3 T_m \cdot D_{imt} + f(N_m) + f(N_m \cdot T_m) + e_{imt} \tag{4.2}$$

We instrument T_m and $T_m \cdot D_{imt}$ with E_m and $E_m \cdot D_{imt}$, where E_m is an indicator for the mother's PANES eligibility, that is, $E_m = 1(N_m > 0)$. Our parameter of interest is β_3 and it measures the average difference in outcomes among children born in eligible and ineligible households across the pre-program. We comment on the validity of this strategy below.

The localized difference in differences approach is valid if: (i) the regression discontinuity identifying assumptions are satisfied, (ii) the difference in differences identifying assumptions are satisfied and (iii) in expectation, the Local Average Treatment Effect (LATE) of treatment effect is the same in the pre-program and program period (Jackson, 2019). In Section 4.4 we showed evidence that (i) is likely satisfied. In addition, in Appendix Table 4.x we report regressions for outcomes during the entire pre-program period. We find no evidence of significant differences in the incidence of low birth weight during pre-program pregnancies. This evidence argues against systematic sorting around the discontinuity. Below we discuss (ii) and (iii).

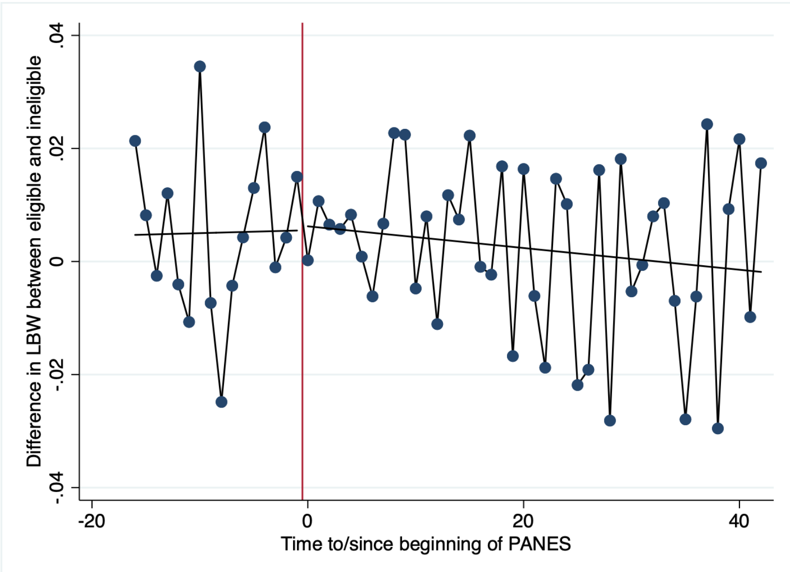
Table 4.x: 2SLS Estimates of the effect of the PANES transfer on low birth weight (<2.500 kg) pre-program data

	(1)	(2)	(3)	(4)	(5)	(6)
1. No Controls	0.003	-0.008	-0.000	-0.025	-0.008	-0.040
	(0.008)	(0.013)	(0.014)	(0.021)	(0.016)	(0.025)
Observations	27835	27835	12102	12102	9362	9362
2. Controls	-0.002	-0.005	-0.004	-0.022	-0.009	-0.035
	(0.009)	(0.013)	(0.013)	(0.021)	(0.015)	(0.025)
Observations	27835	27835	12102	12102	9362	9362
Range	All	All	0.1	0.1	0.075	0.075
Order of polynomial	1	2	1	2	1	2

Note: Each cell corresponds to a different regression. Sample includes pre-program data only. In Columns (1)-(6) we estimate Equation 4.2 for three different ranges around the eligibility threshold and two different orders of polynomial. All estimations include month times year of birth fixed effects, and month times year of baseline visit fixed effects. Row 1 presents regressions with no additional controls while row 2 reports results with the following additional controls: gender of the child, an indicator for whether the mother completed primary school, indicators for geographic department of the household at baseline, for whether the household has centralized hot water, heater, kitchen, microwave, refrigerator, freezer, washing machine, dishwasher, TV, VCR, cable TV, computer, car, whether the block has electricity, piped water, sewage, trash collection, paved streets, sidewalk, whether the home is a house, is owned, and indicators for material of the floor and walls. Controls are included as indicator variables and include a category for missing observations. Standard errors are reported in parentheses and number of observations are reported below each coefficient. * $p<.1$, ** $p<.05$, *** $p<.01$.

The localized difference in differences is valid if there were no changes in eligible households that coincided with eligibility to the program. One may worry that health at birth (incidence of low birth weight) was already improving (decreasing) among eligible households prior to the program. To assess this, Figure 4.ix plots differences in low birth weight (LBW) outcomes between eligible and ineligible mothers giving birth at different months.[33] The x-axis corresponds to the months to and since the beginning of the PANES program (April 2005). Each dot represents the coefficient of the interaction between treatment status and month of birth. For example, the first dot indicates that in February 2003 the incidence of low birth weight among children born from eligible mothers was 2.1 percentage points higher than among those born from ineligible mothers. The solid blue line shows the trend for the difference in low birth weight between eligible and ineligible children. The trend remains constant in approximately 0.05 and there is no indication of a decreasing trend in the pre-program period. This result supports the claim that there were no other changes in health at birth among eligible households in the pre-program period.

Figure 4.ix: Difference in the incidence of low birth weight between eligible and ineligible households



Note: The horizontal axis represents time to/since the beginning of the PANES program in months. The vertical line corresponds to the beginning of the PANES program. Each dot represents the difference in the incidence of low birth weight between children born in eligible and ineligible households in each month. The two solid lines represent the best fit from a linear regression from each side of the cut-off.

Also, the localized difference in differences estimates represent a causal effect of the PANES program if the effect of β_3 is homogeneous. This means that the effect of receiving the transfer while in-utero should be the same for children born in the pre-program period and for children born in the program period. In our setting,

[33]We consider the entire range of the wealth index because, in a given month, the number of observations is significantly reduced when considering smaller bandwidths.

our treatment groups and our control groups contain the same households, easing concerns of mothers differing systematically across periods. However, the localized difference in differences strategy compares children's outcomes among eligible and ineligible households across the pre-program and program period, two very different periods for Uruguay. In 2002, Uruguay was hit by a severe economic crisis and between 2003 and 2005 the economic situation of the country was very adverse. In the period previous to the program, Uruguay started recovering and households' economic situation improved in general. For example, while GDP per capita was on average 8500 USD between 2003 and 2005, it averaged 9500 USD between 2005 and 2007. The unemployment rate also improved as it decreased from 16.7% in 2003 to 9.4% in 2007. The impact of receiving a cash transfer in a context of a severe economic crisis could be different than the one of receiving a transfer when the country is in a better economic situation. Therefore, it is possible that the homogeneity assumption is not valid in this setting.

Chapter 5

Summary and Conclusions

There exists increasing evidence that early childhood is a window of opportunity that cannot be missed. Investments in the zero to five period can be effective to improve child development and, in consequence, to produce gains in long-run outcomes. However, many questions remain regarding where to invest and how to invest and causal evaluations play an important role in this debate. This thesis consists of three essays on the effectiveness of different early childhood programs that were implemented in Uruguay, a small middle income country in Latin America with large socioeconomic gaps in the zero to five period. The three interventions analyzed are very different: Chapter 2 focuses on a parenting intervention, Chapter 3 studies a preschool reform that was implemented by the Uruguayan government by the mid 90's and Chapter 4 evaluates a cash transfer program targeting very poor households.

Specifically, Chapter 2 evaluates the effect of an e-messaging program on parental investment and commitment. The messages were sent to parents or relevant caregivers of children aged 0-2. Recipients received three messages per week during 24 weeks. Both treated and control families attended a parenting workshop of eight weekly group sessions before receiving the messaging program. The design of the messages incorporated behavioral economics insights. Messages were designed to help parents reorient their attention towards positive parenting goals, simplify parental tasks, and reinforce positive identities. The weekly structure of the messages included reminders, concrete suggestion of activities to carry-out at home and reflective and motivational messages. The intervention was evaluated using a randomized controlled trial in 24 Child and Family Care Centers. More than 500 families participated in the randomization. The empirical analysis uses an extensive set of outcomes which were measured using validated instruments and that were self-collected through a parental questionnaire. Unlike previous literature, the assessment focuses on the quantity and also the quality of parental investment, including parental stress, parental sense of competence, parental knowledge about parenting, and parental sensitivity.

The e-messaging intervention evaluated in Chapter 2 increased the quantity and the quality of parental investment. In particular, treated parents had a higher frequency of parental time investment, and higher values of outreach for social support and reflective capacity. The program had stronger effects over families that experienced a lower negative identity at baseline and more negative shocks, suggesting the program triggered the right channels. Overall, the findings from this study suggest that the combination of e-messages and nudges are a promising cost-effective tool to boost parental investment. Also, it is possible to infer from this study that nudges involving recommendations of simple actions and messages of encouragement may be particularly important to activate changes in parental decision-making. Lastly, the results underscore the value of messages to integrate and put into practice previous education experiences.

Going forward, more initiatives such as the one studied in Chapter 2 should be evaluated. Importantly, we need evaluations from non-experimental settings that confirm the effectiveness of such interventions when implemented at scale. Moreover, future research should concentrate on establishing for how long the effects of e-messaging parenting interventions are sustained and what is the optimal duration of such programs. It may be the case that we need to extend the intervention period or complement an early intervention with other ones in later periods of childhood.

Chapter 3 studies the effect of maternal education on the infant health of the next generation. The analysis uses a preschool reform, implemented in the mid 1990 by the Uruguayan government, that substantially increased the availability of school facilities for 4-year-olds. The program was implemented at a differential timing and intensity in different regions so the variation induced by the construction framework can be exploited using a difference-in-differences strategy. The study combines pregnancy and delivery data with school-level data.

The results of Chapter 3 suggest that additional schooling at the beginning of the school trajectory can have positive impacts on the health at birth of the next generation. There is an improvement in health at birth of the offspring of those women that were more exposed to the schooling reform. In particular, there is a reduction in extreme prematurity among first-borns of exposed mothers. One potential channel of the observed effects is that exposed mothers in my sample are more likely to have more than seven prenatal checkups during their pregnancy. Prenatal checkups can be regarded as an indicator of whether a woman is willing to invest in the pregnancy and is an indicator of other healthy behaviors. The preschool reform could have affected health at birth by influencing the decision of women to have fewer children of higher quality and to have children at older ages. This result is in line with a large literature that documents an association between education and fertility choices of women. The study finds that the expansion of preschool facilities decreased teenage pregnancies. However, even in the absence of changes in fertility, the observed effect on health at birth remains, suggesting that this channel does not explain the findings.

The findings from Chapter 3 highlight the importance of education during early childhood and show that the benefits of preschool education can be transmitted across generations. The evidence in this chapter points to preschool education as a way to reduce the intergenerational transmission of poverty due to poor health at birth.

Chapter 4 analyzes the effect of a cash transfer program on long-term educational results of children. The program was implemented between 2005 and 2007 and targeted the poorest 10% households in Uruguay. Eligible households were determined on the basis of a baseline predicted poverty score: households whose score was above a certain threshold were eligible to receive the transfer. This eligibility rule generates a discontinuity that is exploited using a regression discontinuity design. Making use of a rich dataset that links long-term educational outcomes to early life experiences, the analysis compares educational outcomes of children belonging to households just above and just below the eligibility cutoff. The estimations separately consider the impact of PANES on children that were exposed at different stages of early childhood. Three educational outcomes are used: highest grade attained, delay in educational attainment and dropout.

The results in Chapter 4 show that children exposed to the PANES program in the zero to five age range (and after birth) attain higher levels of educational attainment and have a lower incidence of being lagged in educational attainment. There are no long-term education effects on children that were in their mothers' womb during the program period. Following Heckman's model of dynamic complementarity, one would expect that children that received transfers since the in-utero period should have stronger effects on education than those that received them only after birth. However, our findings show the opposite, the effect of the PANES program on education is driven by children exposed to the program during early childhood (and after birth). Finding no effects for the younger cohort of children while positive (but mild) effects for those children that were born earlier in time could be attributed to the fact that the latter started being exposed to the program when Uruguay was in a worst economic situation. Receiving a cash transfer when the economic environment is more adverse could have a stronger impact than receiving a cash transfer in a more prosperous environment.

All in all, the chapters from this thesis suggest that investing in early child development is a good strategy for building human capital and reducing inequities. The three interventions analyzed are very different but the essays in this thesis show that these programs are effective to improve certain short-term and long-term outcomes. From a broader policy perspective, programs that use multiple delivery platforms such as preschool education, parenting programs and cash transfers might be necessary to ensure the highest returns. An important question for future research is to try to understand how best to combine interventions such as the ones studied in this thesis and what is the relative contribution of each component in an effort to find the most cost-effective model.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Abidin, R. (1995). Parenting stress index third edition: Professional manual. lutz, fl: Psychological assessment resources.

Agee, M. D. and Crocker, T. D. (1996). Parental altruism and child lead exposure: Inferences from the demand for chelation therapy. *Journal of Human Resources*, pages 677–691.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Al, C. M., Stams, G. J. J., Asscher, J. J., and van der Laan, P. H. (2014). A programme evaluation of the f amily c risis i ntervention p rogram (fcip): relating programme characteristics to change. *Child & Family Social Work*, 19(2):225–236.

Almond, D., Chay, K. Y., and Lee, D. S. (2005). The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083.

Almond, D. and Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3):153–72.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Amarante, V., Arim, R., and Vigorito, A. (2005). Metodología para la selección de participantes en el plan de emergencia social. *Instituto de Economía, Facultad de Ciencias Económicas, Universidad de la República, Montevideo*.

Amarante, V., Ferrando, M., and Vigorito, A. (2013). Teenage school attendance and cash transfers: An impact evaluation of panes. *Economía*, 14(1):61–96.

Amarante, V., Manacorda, M., Miguel, E., and Vigorito, A. (2016). Do cash transfers improve birth outcomes? evidence from matched vital statistics, program, and social security data. *American Economic Journal: Economic Policy*, 8(2):1–43.

Amarante, V., Manacorda, M., Vigorito, A., and Zerpa, M. (2011). Social assistance and labor market outcomes: Evidence from the uruguayan panes. *Washington, DC: Inter-American Development Bank*.

Amarante, V. and Vigorito, A. (2010). Ccts, social capital and empowerment. evidence from the uruguayan panes. *Unpublished working paper. Universidad de la República*.

ANEP (2005). Panorama de la educación en el uruguay una década de transformaciones 1992–2004. *Montevideo: ANEP*.

ANEP (2007). ¿cuán lejos se está de la universalización de la educación inicial?

ANEP-CODICEN (2000). Una visión integral del proceso de reforma educativa en uruguay, 1995–1999. *Montevideo: ANEP*.

Arranz Freijo, E. B., Artetxe, F. O., Fernández, A. M., Ayala, J. L. M., and Pérez, N. G. (2012). Escala etxadi gangoiti: una propuesta para evaluar los contextos familiares de niños de 2 años de edad1 the etxadi gangoiti scale: a proposal for evaluating the family contexts of two-year-old children. *Revista de Educación*, 358:218–237.

Attanasio, O., Baker-Henningham, H., Bernal, R., Meghir, C., Pineda, D., and Rubio-Codina, M. (2018). Early stimulation and nutrition: the impacts of a scalable intervention. Technical report, National Bureau of Economic Research.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *Bmj*, 349:g5785.

Bailey, D., Duncan, G. J., Odgers, C. L., and Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, 10(1):7–39.

Baird, S., Bohren, J. A., McIntosh, C., and Özler, B. (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, 100(5):844–860.

Baird, S., Ferreira, F. H., Özler, B., and Woolcock, M. (2013). Relative effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries: a systematic review. *Campbell Systematic Reviews*, 9(8).

Baird, S., McIntosh, C., and Özler, B. (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4):1709–1753.

Baker-Henningham, H. and López Bóo, F. (2010). Early childhood stimulation interventions in developing countries: a comprehensive literature review iza discussion paper no. 5282. *Bonn: Institute of Labor Economics.*

Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., and Juffer, F. (2003). Less is more: meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, 129(2):195.

Balsa, A., Gomez, E., Valdes, R., Gonzalez, M., Bloomfield, J., and Cid, A. (2020). Effects of the crianza positiva workshop on parental involvement, parenting knowledge and attitudes. *Working Paper.*

Balsa, A., López Boo, F., Bloomfield, J., Cristiá, A., Cid, A., Ferro, M. d. l. P., Valdés, R., and González, M. d. L. (2021). Effect of crianza positiva e-messaging program on adult-child language interactions. *Behavioral Public Policy*, 1(37).

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind.* MIT press.

Becker, G. S. (1960). An economic analysis of fertility, demographic and economic change in developed countries: a conference of the universities. *National Bureau Commitee for Economic Research*, 209.

Behrman, J. R. and Rosenzweig, M. R. (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92(1):323–334.

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

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Berlinski, S. and Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3):665–680.

Berlinski, S., Galiani, S., and Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92(5-6):1416–1440.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439.

Blattman, C., Fiala, N., and Martinez, S. (2020). The long-term impacts of grants on poverty: Nine-year evidence from uganda's youth opportunities program. *American Economic Review: Insights*, 2(3):287–304.

Blok, H., Fukkink, R., Gebhardt, E., and Leseman, P. (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. *International Journal of Behavioral Development*, 29(1):35–47.

Bloomfield, J. (2020). The effect of maternal education on infant health: Evidence from an expansion of preschool facilities. *Working Paper*.

Bloomfield, J., Balsa, A., and Cid, A. (2021). Using behavioral insights in early childhood interventions: the effects of crianza positiva e-messaging program on parental investment. *Conditionally accepted in Review of Economics of the Household*.

Bloomfield, J. and Cabrera, J. M. (2020). Long-term impacts on education of receiving a cash transfer during early-life. *Working Paper*.

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., and García Coll, C. (2001). The home environments of children in the united states part i: Variations by age, ethnicity, and poverty status. *Child Development*, 72(6):1844–1867.

Breierova, L. and Duflo, E. (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? Technical report, National Bureau of Economic Research.

Bronfenbrenner, U. (1994). Ecological models of human development. *Readings on the Development of Children*, 2(1):37–43.

Bryan, G., Karlan, D., and Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, 2(1):671–698.

Buser, T., Oosterbeek, H., Plug, E., Ponce, J., and Rosero, J. (2017). The impact of positive and negative income changes on the height and weight of young children. *The World Bank Economic Review*, 31(3):786–808.

Cabrera, N. J., Shannon, J. D., Vogel, C., Tamis-LeMonda, C., Ryan, R. M., Brooks-Gunn, J., Raikes, H., and Cohen, R. (2004). Low-income fathers' involvement in their toddlers' lives: Biological fathers from the early head start research and evaluation study. *Fathering: A Journal of Theory, Research & Practice about Men as Fathers*, 2(1).

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24(2):365–389.

Chong, A., Cohen, I., Field, E., Nakasone, E., and Torero, M. (2016). Iron deficiency and schooling attainment in peru. *American Economic Journal: Applied Economics*, 8(4):222–55.

Chou, S.-Y., Liu, J.-T., Grossman, M., and Joyce, T. (2010). Parental education and child health: evidence from a natural experiment in taiwan. *American Economic Journal: Applied Economics*, 2(1):33–61.

Clarke, D., Romano, J. P., and Wolf, M. (2020). The romano–wolf multiple-hypothesis correction in stata. *The Stata Journal*, 20(4):812–843.

Congdon, W. J., Kling, J. R., and Mullainathan, S. (2011). *Policy and Choice: Public Finance through the Lens of Behavioral Economics*. Brookings Institution Press.

Cooklin, A., Giallo, R., and Rose, N. (2012). Parental fatigue and parenting practices during early childhood: An australian community survey. *Child: Care, Health and Development*, 38(5):654–664.

Cortes, K. E., Fricke, H., Loeb, S., and Song, D. S. (2018). Too little or too much? actionable advice in an early-childhood text messaging experiment. Technical report, National Bureau of Economic Research.

Cunha, F. (2015). Subjective rationality, parenting styles, and investments in children. In *Families in an Era of Increasing Inequality*, pages 83–94. Springer.

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

Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4):738–782.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1:697–812.

Cunha, N., Lichand, G., Madeira, R., and Bettinger, E. (2017). What is it about communicating with parents. *Manuscrito não publicado*, pages 1–94.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2):213–238.

Currie, J. and Almond, D. (2011). Human capital development before age five. In *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Damgaard, M. T. and Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, 157:15–26.

De Chaisemartin, C. and D'Haultfœuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. *Available at SSRN 3731856*.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

De Chaisemartin, C. and D'Haultfoeuille, X. (2022). Difference-in-differences estimators of intertemporal treatment effects. Technical report, National Bureau of Economic Research.

de Chaisemartin, C., D'Haultfoeuille, X., and Guyonvarch, Y. (2019). Did_multiplegt: Stata module to estimate sharp difference-in-difference designs with multiple groups and periods.

De Chaisemartin, C. and d'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

Dinçer, M. A., Kaushal, N., and Grossman, M. (2014). Women's education: Harbinger of another spring? evidence from a natural experiment in turkey. *World Development*, 64:243–258.

Doss, C., Fahle, E. M., Loeb, S., and York, B. N. (2017). Supporting parenting through differentiated and personalized text-messaging: Testing effects on learning during kindergarten. cepa working paper no. 16-18. *Stanford Center for Education Policy Analysis*.

Doyle, O., Harmon, C. P., Heckman, J. J., and Tremblay, R. E. (2009). Investing in early human development: timing and economic efficiency. *Economics & Human Biology*, 7(1):1–6.

Doyle, O., Harmon, C. P., and Walker, I. (2005). The impact of parental income and education on the health of their children.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

D'almeida, F. and Margot, D. (2018). La evolución de las telecomunicaciones móviles.

Escueta, M., Quan, V., Nickow, A. J., and Oreopoulos, P. (2017). Education technology: an evidence-based review. Technical report, National Bureau of Economic Research.

Evans, W. D., Wallace, J. L., and Snider, J. (2012). Pilot evaluation of the text4baby mobile health program. *BMC public health*, 12(1):1031.

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

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

Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal*, 12(3):201–218.

Fricke, H., Kalogrides, D., and Loeb, S. (2018). It's too annoying: Who drops out of educational text messaging programs and why. *Economics Letters*, 173:39–43.

Fryer Jr, R. G., Levitt, S. D., and List, J. A. (2015). Parental incentives and early childhood achievement: a field experiment in chicago heights. Technical report, National Bureau of Economic Research.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-term effects of head start. *American Economic Review*, 92(4):999–1012.

Gennetian, L., Darling, M., and Aber, J. L. (2016). Behavioral economics and developmental science: A new framework to support early childhood interventions. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 7(2):2.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, 344(6187):998–1001.

Giné, X., Karlan, D., and Zinman, J. (2010). Put your money where your butt is: a commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, 2(4):213–35.

Gómez-Muzzio, E. and Muñoz-Quinteros, M. (2014). Escala de parentalidad positiva e2p: Manual. *Fundación Ideas para la Infancia*.

Goyette-Ewing, M., Slade, A., Knoebber, K., Gilliam, W., Truman, S., and Mayes, L. (2003). Parents first: a developmental parenting program. yale child study center. *Unpublished Manuscript*.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., and Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the jamaican study. *The Lancet*, 338(8758):1–5.

Grembi, V., Nannicini, T., and Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

Grossman, M. (2006). Education and nonmarket outcomes. *Handbook of the Economics of Education*, 1:577–633.

Güneş, P. M. (2015). The role of maternal education in child health: Evidence from a compulsory schooling law. *Economics of Education Review*, 47:1–16.

Guryan, J., Hurst, E., and Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, 22(3):23–46.

Hack, M. and Fanaroff, A. A. (1999). Outcomes of children of extremely low birthweight and gestational age in the 1990's. *Early human development*, 53(3):193–218.

Hack, M., Flannery, D. J., Schluchter, M., Cartar, L., Borawski, E., and Klein, N. (2002). Outcomes in young adulthood for very-low-birth-weight infants. *New England Journal of Medicine*, 346(3):149–157.

Hack, M., Taylor, H. G., Klein, N., Eiben, R., Schatschneider, C., and Mercuri-Minich, N. (1994). School-age outcomes in children with birth weights under 750 g. *New England Journal of Medicine*, 331(12):753–759.

Hall, C. M. and Bierman, K. L. (2015). Technology-assisted interventions for parents of young children: Emerging practices, current research, and future directions. *Early childhood research quarterly*, 33:21–32.

Hart, B. and Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Paul H Brookes Publishing.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.

Haushofer, J. and Shapiro, J. (2018). The long-term impact of unconditional cash transfers: experimental evidence from kenya. *Busara Center for Behavioral Economics, Nairobi, Kenya.*

Havnes, T. and Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.

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

Heckman, J. J., Krueger, A. B., et al. (2005). Inequality in america: What role for human capital policies? *MIT Press Books*, 1.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of Public Economics*, 94(1-2):114–128.

Heß, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal*, 17(3):630–651.

Holmlund, H., Lindahl, M., and Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, 49(3):615–51.

Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., and Wartella, E. (2015). Supporting head start parents: impact of a text message intervention on parent–child activity engagement. *Early Child Development and Care*, 185(9):1373–1389.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Jackson, C. K. (2019). Can introducing single-sex education into low-performing schools improve academics, arrests, and teen motherhood? *Journal of Human Resources*, pages 0618–9558R2.

Johnston, C. and Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology*, 18(2):167–175.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, 62(12):3393–3411.

Kirby, K. N., Petry, N. M., and Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: general*, 128(1):78.

Knerr, W., Gardner, F., and Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low-and middle-income countries: A systematic review. *Prevention Science*, 14(4):352–363.

Koch, A., Nafziger, J., and Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115:3–17.

Lavecchia, A. M., Liu, H., and Oreopoulos, P. (2016). Behavioral economics of education: Progress and possibilities. In *Handbook of the Economics of Education*, volume 5, pages 1–74. Elsevier.

Lawrance, E. C. (1991). Poverty and the rate of time preference: evidence from panel data. *Journal of Political Economy*, 99(1):54–77.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Leer, J., López Bóo, F., Perez Expósito, A., and Powell, C. (2016). A snapshot on the quality of seven home visit parenting programs in latin america and the caribbean.

Lindeboom, M., Llena-Nozal, A., and van der Klaauw, B. (2009). Parental education and child health: Evidence from a schooling reform. *Journal of Health Economics*, 28(1):109–131.

Llambí, C., Perera, M., and Messina, P. (2009). Desigualdad de oportunidades y el rol del sistema educativo en los logros de los jóvenes uruguayos. *Documento de trabajo del CINVE*, (04).

López Bóo, F., Cubides Mateus, M., Sorio, R., Garibotto, G., and Berón, C. (2019). Measuring the quality of the home environment of young children in uruguay: Socioeconomic gradients in the home inventory.

Madrian, B. C. and Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, 116(4):1149–1187.

Main, M. (1991). Metacognitive knowledge, metacognitive monitoring, and singular (coherent) vs. multiple (incoherent) models of attachment. *Attachment Across the Life Cycle*, 127:159.

Manacorda, M., Miguel, E., and Vigorito, A. (2011). Government transfers and political support. *American Economic Journal: Applied Economics*, 3(3):1–28.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *Science*, 341(6149):976–980.

Marroig, A., Perazzo, I., Salas, G., and Vigorito, A. (2017). Evaluación de impacto del programa de acompañamiento territorial de uruguay crece contigo.

Martin, E., Weiland, C., and Page, L. C. (2018). Text-based mentoring for postpartum mothers: A feasibility study. *Early Child Development and Care*, pages 1–24.

Mayer, S. E., Kalil, A., Oreopoulos, P., and Gallegos, S. (2018). Using behavioral insights to increase parental engagement: The parents and children together intervention. *Journal of Human Resources*, pages 0617–8835R.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

McCrary, J. and Royer, H. (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review*, 101(1):158–95.

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, 64(1):95–114.

Menéndez, S., Jiménez, L., and Hidalgo, M. V. (2011). Estructura factorial de la escala psoc (parental sense of competence) en una muestra de madres usuarias de servicios de preservación familiar. *Revista Iberoamericana de Diagnóstico y Evaluación e Avaliação Psicológica*, 2(32):187–204.

Meuwissen, A., Giovanelli, A., Labella, M., and Susman-Stillman, A. (2017). Text2learn: An early literacy texting intervention by community organizations.

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., and Stampini, M. (2019). Long-term impacts of conditional cash transfers: review of the evidence. *The World Bank Research Observer*, 34(1):119–159.

Mischel, W., Ebbesen, E. B., and Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, 21(2):204.

National Institute of Child Health and Human Development Early Child Care Research Network (2006). Child-care effect sizes for the NICHD study of early child care and youth development. *American Psychologist*, 61(2):99–116.

Nores, M. and Barnett, W. S. (2010). Benefits of early childhood interventions across the world:(under) investing in the very young. *Economics of Education Review*, 29(2):271–282.

Oreopoulos, P., Stabile, M., Walld, R., and Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. *Journal of Human Resources*, 43(1):88–138.

Pabilonia, S. W. and Song, Y. (2013). Single mothers' time preference, smoking, and enriching childcare: Evidence from time diaries. *Eastern Economic Journal*, 39(2):227–255.

Patrick, K., Raab, F., Adams, M., Dillon, L., Zabinski, M., Rock, C., Griswold, W., and Norman, G. (2009). A text message-based intervention for weight loss: randomized controlled trial. *Journal of Medical Internet Research*, 11(1):e1.

Pickering, J. A., Sanders, M. R., et al. (2014). The triple p-positive parenting program: An example of a public health approach to evidence-based parenting support. *Family Matters*, (96):53.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, 36(1):11–46.

Romano, J. P. and Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108.

Romano, J. P. and Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

Romano, J. P. and Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, 113:38–40.

Rosenzweig, M. R. and Zhang, J. (2013). Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of chinese twins. *Journal of Development Economics*, 104:245–260.

Rosero, J. and Oosterbeek, H. (2011). Trade-offs between different early childhood interventions: Evidence from ecuador.

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

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N., and Grantham-McGregor, S. (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in bogota. *Journal of Human Resources*, 50(2):464–483.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59.

Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., Lopez-Boo, F., Macours, K., Marshall, D., Paxson, C., et al. (2015a). Wealth gradients in early childhood cognitive development in five latin american countries. *Journal of Human Resources*, 50(2):446–463.

Schady, N. et al. (2015b). *The early years: child well-being and the role of public policy*. Springer.

Schilbach, F., Schofield, H., and Mullainathan, S. (2016). The psychological lives of the poor. *American Economic Review*, 106(5):435–40.

Shah, A. K., Mullainathan, S., and Shafir, E. (2012). Some consequences of having too little. *Science*, 338(6107):682–685.

Slade, A. (1999). Attachment theory and research: Implications for the theory and practice of individual psychotherapy with adults.

Strauss, J. and Thomas, D. (1995). Human resources: Empirical modeling of household and family decisions. *Handbook of Development Economics*, 3:1883–2023.

Tansini, R. (2008). Resultados escolares en escuelas públicas de montevideo:¿ de qué dependen? *Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República, Uruguay.*

Terra, J. P. (1988). Pobreza y ninez en el uruguay actual. *Cuadernos del CLAEH*, 47:47–68.

Thaler, R. H. and Ganser, L. (2015). *Misbehaving: The making of behavioral economics*. WW Norton New York.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40(4):1008–1018.

Waldfogel, J. and Washbrook, E. (2011). Early years policy. *Child Development Research*, 2011.

Walker, S. P., Chang, S. M., Vera-Hernández, M., and Grantham-McGregor, S. (2011). Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics*, 127(5):849–857.

Walker, S. P., Powell, C., Chang, S. M., Baker-Henningham, H., Grantham-McGregor, S., Vera-Hernandez, M., and López-Boo, F. (2015). Delivering parenting interventions through health services in the caribbean: impact, acceptability and costs. Technical report, IDB Working Paper Series.

Wilson, M. H. (2010). *Behavioral management parent training for children with developmental delays ages 18-36 months*. Washington State University.

Yeung, W. J., Linver, M. R., and Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, 73(6):1861–1879.

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

Zeanah, C. H. (2000). Disturbances of attachment in young children adopted from institutions. *Journal of Developmental and Behavioral Pediatrics*, 21(3):230–236.

Zeger, S. L. and Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, pages 121–130.

Samenvatting (Summary in Dutch)[1]

De laatste jaren zijn er steeds meer aanwijzingen dat de omstandigheden in het vroege leven gevolgen kunnen hebben voor de ontwikkeling en het welzijn van mensen op lange termijn. Investeringen wanneer kinderen tussen nul en vijf jaar oud zijn, kunnen de toekomstige uitkomsten effectief beïnvloeden. Zulke investeringen zijn belangrijk omdat zij de situatie van kinderen die in armoede opgroeien, kunnen verbeteren. Beleidsmakers over de hele wereld voeren een breed scala aan beleidsmaatregelen voor jonge kinderen uit. Dit betreft programma's voor financiële steun, de uitbreiding van centra voor jonge kinderen en voorschoolse opvang, en interventieprogramma's met ouders. De vraag hoe maatregelen voor jonge kinderen het best kunnen worden uitgevoerd, is echter niet eenvoudig te beantwoorden. Causale evaluaties van beleid voor jonge kinderen spelen een belangrijke rol in het debat over waar te investeren en hoe het best te investeren.

Dit proefschrift bestaat uit drie essays over de effectiviteit van verschillende programma's voor jonge kinderen die werden uitgevoerd in Uruguay, een land met een gemiddeld inkomen in Latijns-Amerika met grote sociaal-economische verschillen in de vroege kinderjaren.

Hoofdstuk 2 evalueert het effect van een e-messaging-programma op de investering en betrokkenheid van ouders. Berichten werden verstuurd naar ouders of verzorgers van kinderen van 0-2 jaar. Meer dan 500 gezinnen van 24 centra voor kinder- en gezinszorg namen deel aan de studie. De gezinnen werden willekeurig verdeeld in een behandelgroep, die gedurende 24 weken drie keer per week berichten ontving, en een controlegroep. Zowel de behandelde als de controlegezinnen namen deel aan een workshop van acht wekelijkse groepssessies voordat zij het berichtenprogramma ontvingen. Het ontwerp van de berichten was gebaseerd op gedragseconomische inzichten. De berichten werden ontworpen om ouders te helpen hun aandacht te heroriënteren naar positieve opvoedingsdoelen, ouderlijke taken te vereenvoudigen en positieve identiteiten te versterken. De wekelijkse structuur van de berichten omvatte herinneringen, concrete suggesties voor thuis uit te voeren activiteiten en reflectieve en motiverende boodschappen. De interventie werd geëvalueerd aan de hand van gevalideerde instrumenten die zelf werden verzameld via een oudervra-

[1] Deze samenvatting is tot stand gekomen met behulp van Hessel Oosterbeek.

genlijst. In tegenstelling tot eerdere literatuur richt de beoordeling zich zowel op de kwantiteit als op de kwaliteit van de ouderlijke investering, waaronder ouderlijke stress, ouderlijk gevoel van competentie, ouderlijke kennis over ouderschap en ouderlijke sensitiviteit.

De in hoofdstuk 2 geëvalueerde e-messaging interventie verhoogde de kwantiteit en de kwaliteit van de ouderlijke investering. In het bijzonder hadden de behandelde ouders een hogere frequentie van ouderlijke tijdsinvestering, en hogere waarden van bereik voor sociale steun en reflectief vermogen. Het programma had sterkere effecten op gezinnen die bij aanvang een meer negatieve identiteit en meer negatieve schokken ervoeren. Dit wijst erop dat het programma de juiste kanalen activeerde. In het algemeen suggereren de bevindingen van deze studie dat de combinatie van e-berichten en nudges veelbelovende kosteneffectieve instrumenten zijn om ouderlijke investeringen te stimuleren. Ook kan uit deze studie worden afgeleid dat nudges met aanbevelingen voor eenvoudige acties en berichten van aanmoediging bijzonder belangrijk kunnen zijn om veranderingen in de besluitvorming van ouders te activeren. Ten slotte onderstrepen de resultaten de waarde van boodschappen om eerdere onderwijservaringen te integreren en in praktijk te brengen.

Hoofdstuk 3 schat het effect van het op jongere leeftijd naar school gaan van meisjes op de gezondheid bij de geboorte van hun kinderen. Ik evalueer de gezondheid bij de geboorte van de nakomelingen van moeders die op 4-jarige leeftijd aan een hervorming werden blootgesteld. De identificatiestrategie maakt gebruik van een bouwprogramma voor openbare voorschoolse voorzieningen dat midden jaren negentig in Uruguay werd uitgevoerd en waardoor de beschikbaarheid van voorschoolse voorzieningen aanzienlijk toenam. Ik benut de variatie in het aantal gebouwde voorzieningen tussen de regio's en in de tijd. De belangrijkste databrom die in de analyse wordt gebruikt, combineert informatie van geboortegegevens met gegevens op schoolniveau.

De resultaten van hoofdstuk 3 geven aan dat extra scholing aan het begin van het schooltraject positieve gevolgen kan hebben voor de gezondheid bij de geboorte van de volgende generatie. Er is een verbetering van de gezondheid bij de geboorte van de nakomelingen van die vrouwen die werden blootgesteld aan de schoolhervorming. Een mogelijk kanaal voor de waargenomen effecten is dat blootgestelde moeders in mijn steekproef vaker meer dan zeven prenatale controles ondergaan tijdens hun zwangerschap. Prenatale controles kunnen worden beschouwd als een indicator dat een vrouw bereid is te investeren in de zwangerschap en als indicator voor ander gezond gedrag. Bovendien stel ik vast dat de uitbreiding van het aantal plaatsen in het kleuteronderwijs een effect heeft gehad op de vruchtbaarheid. Dat is vooral te danken is aan een vermindering van het aantal tienerzwangerschappen.

De bevindingen van hoofdstuk 3 wijzen op het belang van onderwijs in de vroege kinderjaren en laten zien dat de voordelen van voorschools onderwijs kunnen worden overgedragen op andere generaties. De gegevens in dit hoofdstuk wijzen erop dat voorschools onderwijs een manier is om de intergenerationele overdracht van armoede als gevolg van een slechte gezondheid bij de geboorte te verminderen.

Hoofdstuk 4 analyseert het effect van financiële steun aan gezinnen op de onderwijsresultaten van kinderen op lange termijn. Het steunprogramma werd uitgevoerd tussen 2005 en 2007 en was gericht op de armste 10 procent huishoudens in Uruguay. De in aanmerking komende huishoudens werden bepaald op basis van een armoedeindex. Huishoudens met een score boven een bepaalde drempel kwamen in aanmerking voor de overdracht. Door gebruik te maken van een rijke dataset die de onderwijsresultaten op lange termijn koppelt aan vroege levenservaringen, vergelijkt de analyse de onderwijsresultaten van kinderen die behoren tot huishoudens net boven en net onder de drempel om in aanmerking te komen. De schattingen houden afzonderlijk rekening met het effect van financiële steun op kinderen die in verschillende stadia van hun vroege jeugd zijn blootgesteld.

De resultaten in hoofdstuk 4 laten zien dat kinderen die tussen nul en vijf jaar (en na de geboorte) aan het steunprogramma zijn blootgesteld, een hoger onderwijsniveau bereiken en minder vaak een onderwijsachterstand hebben. Er zijn geen onderwijseffecten op lange termijn voor kinderen die tijdens de programmaperiode nog in de buik van hun moeder zaten.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. For a full list of PhD theses that appeared in the series we refer to List of PhD Theses – Tinbergen.nl. The following books recently appeared in the Tinbergen Institute Research Series:

754 A.C. VAN VLODROP, *Essays on Modeling Time-Varying Parameters*
755 J. SUN, *Tell Me How To Vote, Understanding the Role of Media in Modern Elections*
756 J.H. THIEL, *Competition, Dynamic Pricing and Advice in Frictional Markets: Theory and Evidence from the Dutch Market for Mortgages*
757 A. NEGRIU, *On the Economics of Institutions and Technology: a Computational Approach*
758 F. GRESNIGT, *Identifying and Predicting Financial Earth Quakes using Hawkes Processes*
759 A. EMIRMAHMUTOGLU, *Misperceptions of Uncertainty and Their Applications to Prevention*
760 A. RUSU, *Essays in Public Economics*
761 M.A. COTOFAN, *Essays in Applied Microeconomics: Non-Monetary Incentives, Skill Formation, and Work Preferences*
762 B.P.J. ANDRÉE, *Theory and Application of Dynamic Spatial Time Series Models*
763, P. PELZL, *Macro Questions, Micro Data: The Effects of External Shocks on Firms*
764 D.M. KUNST *Essays on Technological Change, Skill Premia and Development*
765 A.J. HUMMEL, *Tax Policy in Imperfect Labor Markets*
766 T. KLEIN, *Essays in Competition Economics*
767 M. VIGH, *Climbing the Socioeconomic Ladder: Essays on Sanitation and Schooling*
768 YAN XU, *Eliciting Preferences and Private Information: Tell Me What You Like and What You Think*
769 S. RELLSTAB, *Balancing Paid Work and Unpaid Care over the Life-Cycle*
770 Z. DENG, *Empirical Studies in Health and Development Economics*
771 L. KONG, *Identification Robust Testing in Linear Factor Models*
772 I. NEAMŢU, *Unintended Consequences of Post-Crisis Banking Reforms*
773 B. KLEIN TEESELINK, *From Mice to Men: Field Studies in Behavioral Eco-*

nomics

774 B. TEREICK, *Making Crowds Wiser: The Role of Incentives, Individual Biases, and Improved Aggregation*

775 A. CASTELEIN, *Models for Individual Responses*

776 D. KOLESNYK, *Consumer Disclosures on Social Media Platforms: A Global Investigation*

777 M.A. ROLA-JANICKA, *Essays on Financial Instability and Political Economy of Regulation*

778 J.J. KLINGEN, *Natural Experiments in Environmental and Transport Economics*

779 E.M. ARTMANN, *Educational Choices and Family Outcomes*

780 F.J. OSTERMEIJER, *Economic Analyses of Cars in the City*

781 T. ÖZDEN, *Adaptive Learning and Monetary Policy in DSGE Models*

782 D. WANG, *Empirical Studies in Financial Stability and Natural Capital*

783 L.S. STEPHAN, *Estimating Diffusion and Adoption Parameters in Networks New Estimation Approaches for the Latent-Diffusion-Observed-Adoption Model*

784 S.R. MAYER, *Essays in Financial Economics*

785 A.R.S. WOERNER, *Behavioral and Financial Change – Essays in Market Design*

786 M. WIEGAND, *Essays in Development Economics*

787 L.M. TREUREN, *Essays in Industrial Economics - Labor market imperfections, cartel stability, and public interest cartels*

788 D.K. BRANDS, *Economic Policies and Mobility Behaviour*

789 H.T.T. NGUYN, *Words Matter? Gender Disparities in Speeches, Evaluation and Competitive Performance*

790 C.A.P BURIK, *The Genetic Lottery. Essays on Genetics, Income, and Inequality*

791 S.W.J. OLIJSLAGERS, *The Economics of Climate Change: on the Role of Risk and Preferences*

792 C.W.A. VAN DER KRAATS, *On Inequalities in Well-Being and Human Capital Formation*

793 Y. YUE, *Essays on Risk and Econometrics*

794 E.F. JANSSENS, *Estimation and Identification in Macroeconomic Models with Incomplete Markets*

795 P.B. KASTELEIN, *Essays in Household Finance: Pension Funding, Housing and Consumption over the Life Cycle*

796 J.O. OORSCHOT, *Extremes in Statistics and Econometrics*

797 S.D.T. HOEY, *Economics on Ice: Research on Peer Effects, Rehiring Decisions and Worker Absenteeism*

798 J. VIDIELLA-MARTIN, *Levelling the Playing Field: Inequalities in early life conditions and policy responses*

799 Y. XIAO, *Fertility, parental investments and intergenerational mobility*

800 X. YU, *Decision Making under Different Circumstances: Uncertainty, Urgency, and Health Threat*

801 G. GIANLUCA, *Productivity and Strategies of Multiproduct Firms*

802 H. KWEON, *Biological perspective of socioeconomic inequality*

803 D.K. DIMITROV, *Three Essays on the Optimal Allocation of Risk with Illiquidity, Intergenerational Sharing and Systemic Institutions*

This thesis investigates the effectiveness of different early childhood interventions that aim to help families either by providing parental tools to improve parenting practices and overcome behavioral biases, expanding access to local educational resources, and providing cash to overcome financial constraints. The first essay studies whether an e-messaging program rooted on behavioral economics insights is effective to increase parental investment and reinforce parental commitment. The second essay estimates the effect of girls starting school earlier on health at birth of the next generation. The final essay evaluates the long-term effects of receiving unconditional cash transfers since conception and up to the first five years of life on education outcomes.

Juanita Bloomfield studied Economics at Universidad de Montevideo. She continued her studies at the Tinbergen Institute, following the Master of Philosophy in Economics. After finishing the research master she started a PhD degree at the University of Amsterdam. She currently works as an Assistant Professor at the Department of Economics of Universidad de Montevideo.