Estimating Inter-generational Returns to Medical Care: New Evidence from At-Risk Newborns*

Damian Clarke† Nicolás Lillo Bustos‡ Kathya Tapia-Schythe§

June 6, 2024

Abstract

We examine whether intensive early-life interventions targeted to at-risk newborns are transmitted intergenerationally. Using a regression discontinuity design and administrative data from Chile we document the surprising fact that children of individuals who were treated at birth have substantially worse indicators of health at birth a generation later. We suggest this owes to selective fertility, finding that marginally treated individuals are substantially more likely to give birth. These new stylised facts suggest that in certain circumstances, the long-term implications of health policies within family lineages may be quite different to their short-term implications, placing more weight on necessary reinforcing interventions.

JEL Codes: I11; I18; J13; H51; O15.
Keywords: Early life interventions; inter-generational mobility; parental investments; fertility; health care provision.

*We thank Jason Lindo, Anastasia Terskaya, Jorge Rodríguez and Atheen Venkataramani for thoughtful comments and conversations, and Jorge Fabres and Javier Cifuentes for information related to medical interventions and institutional details on neonatal health-care practices. We are grateful to seminar audiences at the Barcelona GSE Summer Forum, IADB, LACEA, Pontificia Universidad Javeriana, University of Chile, University of Exeter, Universidad de la Plata, New School of Economics, and Universidad O'Higgins. We thank Sofía Arredondo, Agustina Crozier, and Daniel Pailañir for outstanding research assistance. Clarke acknowledges financial support from ANID Chile (FONDECYT Regular 1200634) and support from the ANID/Millennium Science Initiative Grant via the "Millennium Institute for Market Imperfections and Public Policy (MIPP)".

†Department of Economics, University of Exeter, Department of Economics, University of Chile, IZA, CAGE & MIPP. Address: Diagonal Paraguay 257, Santiago, Chile. Contact email: dclarke@fen.uchile.cl.
‡Ministry of Economics, Chile. Contact: nlillob@economia.cl.
§Department of Economics, University of California, Davis. Contact: kattapia@ucdavis.edu.

1 Introduction

Returns to early life investment programs accrue over life. How generalized is this accrual, for how long does it last, and over what dimensions do investments accrue? These questions have important implications for policy design related to optimal welfare investments over the life-cycle (Cunha et al., 2010; Currie and Rossin-Slater, 2015). Conceivably, if early-life circumstances shape later life outcomes, policies which affect such investments could have considerable inter-generational implications (Mogstad and Torsvik, 2022), suggesting impacts may echo well beyond the period of policy receipt.

In this paper we seek to understand these questions, studying whether targeted and intensive medical interventions in very low birth weight (VLBW) children immediately after birth are transmitted onto the outcomes of *future* generations. Using a multigenerational linked database of the universe of births in Chile and a regression discontinuity (RD) design, we examine whether women who were born weighing just below a specific (1,500 gram) treatment cut-off go on to have babies who have greater measurable stocks of health at birth a generation later. We trace out the impacts of an individuals' early life treatment on their future interactions with the health care system, the composition of their own family during childhood, their fertility behaviour upon maturity, and ultimately, the inter-generational transmission of intensive medical care receipt to measures of their children's health and well-being.

Large public health programs targeted to children with poor birth outcomes result in immediate improvements in health and survival (Almond et al., 2010; Bharadwaj et al., 2013; Chyn et al., 2021), which impact educational outcomes in childhood and adolescence (Bharadwaj et al., 2013; Daysal et al., 2022), reduce future reliance on public social safety net programs (Chyn et al., 2021) and positively affect other members of the family including siblings and mothers over both the short and long term (Daysal et al., 2022).[1] More generally, there is substantial evidence from a broad literature in economics documenting intergenerational transmission of health at birth (Currie and Moretti, 2007; Giuntella et al., 2019; Lahti-Pulkkinen et al., 2017; Royer, 2009, for example). At the same time, the way individuals interact with health policies, and the degree to which treatment receipt can shape future outcomes across the life course, is complex and multi-faceted. Both Currie and Rossin-Slater (2015) and Wüst (2022), reviewing a broad range of studies, demonstrate that early life policies have appreciable impacts across multiple later life domains including adult health, educa-

[1] These studies, based on a similar RD design as that used in this paper, point to VLBW cut-offs as significant determinants of well-being for individuals during early life and childhood. The study of Bharadwaj et al. (2013) considers the first generation using the same data and the same context studied in this paper. Their results document a strong and enduring 'first-stage' to the longer term and intergenerational context we study in this paper. As we discuss in the methods and results section of this paper, we document that their earlier results still hold when working with a much extended sample and recent advances in optimal RD designs (as has been documented in other recent RD papers of this nature such as Daysal et al. (2022)). We also discuss how our results take on board important critiques relating to heaping of the running variable (birth weight) in this case.

tional attainment, labor market outcomes, and other measures of socio-economic status. If neonatal and early life health programs lift up the life courses of their original recipients, and in turn spillover to future generations, this may suggest that the already large benefits of such programs could be a (considerable) lower bound. However, if families or individuals change their behaviour as a downstream result of medical intervention (or lack of intervention), or if medical intervention shapes intermediate outcomes potentially shifting the life-courses of individuals, the inter-generational returns to medical care may be entirely different to returns to the first generation. Indeed, intergenerational returns could even be negative if families make compensating investments in less treated individuals (as has been observed empirically (Bharadwaj et al., 2018)) thus overwhelming initial investments received by more intensively treated individuals, or if medical receipt shapes future individual behaviours in other unexpected ways.[2]

In this study we trace out impacts of the receipt of intensive medical care very early in childhood on outcomes in childhood, adolescence, early adulthood, and spillovers into future generations. Linking comprehensive microdata registries, we are able to observe all births in Chile between 1992-2018, their future inpatient hospitalizations, and their future fertility histories, and, for those women that go on to have *their own* births, we observe the early life health stocks and survival records for their children. Our results show that despite relatively large effects of the program on infant mortality and days of hospitalization in the first generation suggesting a substantive first generation impact (Bharadwaj et al., 2013), there is clearly no positive inter-generational transmission to the health outcomes at birth of the *children* of individuals who were treated at birth. Indeed, more surprisingly, we observe quite clear negative transmissions of intensive medical treatment receipt at birth to the health outcomes of second generation of children, especially when studying outcomes very low in the distribution of health at birth, such as the likelihood that a child is born prematurely, or with a very low birth weight.

We consider a number of channels which may explain the reversal of impacts across generations. These include selection out of fertility, a reduction in the likelihood of second generation fetal death resulting in negative selection of marginal surviving individuals, and (first generation) parental reinforcing behaviours overwhelming initial positive effects of intensive early life health treatments. We additionally consider the likelihood that the lack of observed effects owes to low statistical power. In conducting these tests, we find clear evidence of strong policy effects on fertility. These results suggest a chain of policy impacts, whereby individuals who receive intensive medical treatment are more able to take births to term, with subsequent

[2]There is evidence, for example, that early life health shocks may interact with future events and investments, potentially magnifying the impact of initial shocks (Duque et al., 2019). In general, family circumstance has large and pervasive intergenerational implications (see for example Eshaghnia et al. (2022)), which in the absence of some exogenous shock, may make pinning down the causal intergenerational impact of early life health-care receipt difficult.

negative implications on average health stocks of affected cohorts of children. The results and stylized facts we document suggest revisions are required to the standard understanding of intergenerational transfers when policy can shape outcomes on both extensive (fertility) and intensive (health at birth) outcomes. These results provide a counter-example to the broad and growing literature of positive inter-generational transfers, at least if only intensive margin outcomes are observed.

This result is novel, given that the overwhelming body of the literature suggests that there are positive intergenerational spillovers in health and in socioeconomic measures more broadly, and that such positive spillovers are observable in mean outcomes of the population. For example, in examining health shocks, Lee (2014) points to intergenerational impacts still persistent in grand-children due to their grandmother's exposure to a violent uprising while pregnant, with similar patterns observed when considering in utero declines in pollution exposure as a result of the US Clean Air Act (Colmer and Voorheis, 2020). In a recent study East et al. (2023) document important positive intergenerational spillovers owing to Medicaid coverage in the prenatal period. More broadly, there is evidence to suggest that maternal exposure to stressful events in adolescence and adulthood is passed across generations (Akresh et al., 2021). Similarly, Almond et al. (2012) document intergenerational transfer of exposure to disease across generations of mothers. Bhalotra and Rawlings (2011, 2013) document significant gradients in the exposure to health shocks of children based on the health of their mothers, additionally pointing to important channels of intergenerational transmission of health. These first-generation exposures have also been shown to be reflected in later life health outcomes of second generation children (Venkataramani, 2011), to be observed in long-term measures such as life expectancy (Black et al., 2022), and to endure across more than two generation (Costa, 2021). A rich stream of emerging literature has shown intergenerational transfers in education and well-being, with important implications on the accrual of inequality over generations.[3]

This paper contributes to a range of literatures. It joins a number of recent studies in documenting intergenerational transmission of early life health indicators. It also contributes to a literature considering the intergenerational transfer of exposures to particular events or environments across cohorts. It additionally joins a large literature considering the returns to public health programmes. This is one of only a small number of studies to consider intergenerational implications of individual-level exposure to a large public health policy, allowing for the estimation of long-term returns of such programmes.[4] The study closest to ours in

[3] Two studies have also considered the interaction of educational and health policies: Rossin-Slater and Wüst (2020) examining nurse visits and pre-school, and Barr and Gibbs (2022) who examine head start which includes health screenings, both finding substantial positive intergenerational effects.

[4] There are a number of studies which consider how exposure to policies can impact educational or labour market of exposed individuals lifting up trajectories of disadvantaged individuals, in this way breaking intergenerational cycles of disadvantage. Examples of such policies, particularly in the Scandinavian context, are discussed by Wüst (2022); see for example Bütikofer et al. (2019).

this regard is East et al. (2023): an impressive multigenerational study of Medicaid, finding positive intergenerational spillovers in the US. Importantly, our study and their study work with a broadly similar age range of second generation mothers (though of course a quite different exposure variable), suggesting a useful contrast to our proof of concept that *in some settings* intergenerational spillovers can be negative when studying observed health outcomes. Perhaps the main contribution of this study, and where it expands knowledge in each of the literatures discussed above, is in documenting that the marginal returns to medical care in one generation can have considerable long-term implications, improving the outcomes of individuals in the first generation, while at the same time bequeathing weaker health stocks to the following generation, at least when considering measures of health at birth.[5]

The structure of the paper is as follows. In Section 2 we discuss the context studied here and the nature of the VLBW assignment threshold as a determinant of early life medical care. Section 3 discusses the intergenerational linked administrative data generated for this study, and Section 4 discusses how this is used to estimate impacts of exposure based on an RD design, additionally presenting a simple model to understand the full content of RD estimates in this context. Section 5 provides results as well as identification checks, discussing both mechanisms which could explain the observed results, and their implications on how we should conceive the (long-term) returns to medical care. Section 6 briefly concludes.

2 Background and Medical Care Regimes

Birth registration in Chile is universal, and the large majority of births (over 99%) are attended in public hospitals or private clinics which follow national-level protocols set by the Ministry of Health. Births are overwhelmingly attended by doctors and/or midwives. As laid out in Bharadwaj et al. (2013), national-level guidelines were set in 1991 by a national committee to standardize treatments at Neonatal Intensive Care Units (NICUs) in the country, which exist in each of Chile's 16 regions. A particular concern at this point was the high rate of infant mortality among very low birth weight infants (weights below 1,500 grams). Treatment protocols often explicitly mention 1,500 grams as a treatment cut-off, and these births undergo a systematic treatment protocol with follow-up procedures (Hübner et al., 2009).

What is different in this paper is the consideration of how exposure to a health policy is transmitted into *future* generations.

[5] This papers also contributes, in a limited way, to a literature on the replicability crisis in social sciences, and concerns with the use of non-public data. For a particular study (Bharadwaj et al., 2013) when previously private data was subsequently published without restrictions, we were able to substantively replicate the findings of the original paper's private data, starting from scratch in data collation and generation. Further, updating results based on technical advances in the intervening period since the paper's publication, and additionally updating to include a substantially longer time-frame which more than duplicated the original sample points to results that are entirely consistent with those in the originally published research.

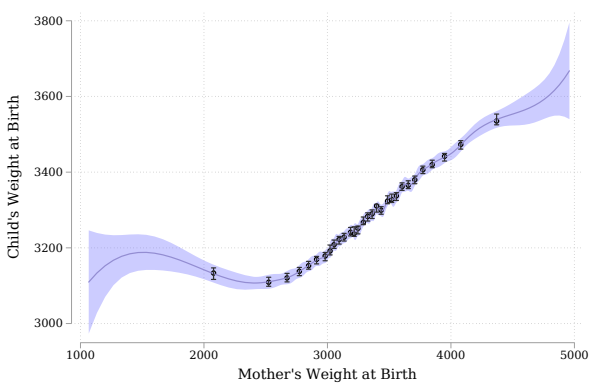
Bharadwaj et al. (2013), pointing to Gonzalez et al. (2006); Jiménez and Romero (2007); Palomino et al. (2005) note a number of explicit treatment cut-offs which are documented at 1,500 grams. This includes the use of artificial lung surfactant, a complementary nutrition program providing specialized supplementation (*PNAC prematuro*), and a health care reform (*AUGE*) defining neonatal follow up appointments to screen for particular pathologies which are targeted at infants born at less than 1,500 grams. These criteria are explicitly recognized also in later policy documents, such as the National neonatal guidebook issued to medical practitioners (Mena Nanning et al., 2005),[6] and when formally included as requirements for accessing policies (such as that indicated in the *AUGE* reform), these require children to have weights *strictly below* 1,500 grams, implying that those who weigh 1,500 grams will not classify for treatment.

Discontinuous treatment assignments at 1,500 grams have been apparent in official national clinical guidelines for an extended period of time. The 2005 Neonatal Guide, referring to actions taken in the previous decade, notes policies targeted to reduce rates of morbidity among infants born with weights of less than 1,500 grams such as the national surfactant program (and other programs to standardise use of drugs such as Indometacin), programs to follow infants over an extended period of time post-birth, the regionalization of neonatal services (Mena Nanning et al., 2005, pp. 15-16), and recommendations that all newborns with a weight less than 1,500 grams should receive diagnostic ultrasounds to examine intracranial haemorrhage, diagnostic tests for retinopathy, yearly specialised opthamological follow-ups, as well as a range of other specialised treatment courses in cases of particular diagnoses. There is evidence that the discontinuity in treatment continues to be highly relevant even late in the period under study. For example, the Clinical Manual for Neonatal Care written in 2016 for one of Santiago's large public hospitals with a high complexity neonatal care unit lists a large number of specific treatments which neonates should receive if they weigh under 1,500 grams. These include specialised procedures for temperature maintenance at birth, a delayed clamping of the umbilical chord to allow for greater blood flow to the baby following birth, the required presence of two specialists trained in re-animation, specific formulae for ventilation and nasal air flow masks, and direct transfer to neonatal intensive care units. An entire chapter is dedicated to these procedures in the hospital's clinical guides (Mulhaüsen Muñoz and González, 2016). No similar such guidelines exist for other distributional points of birth weight.

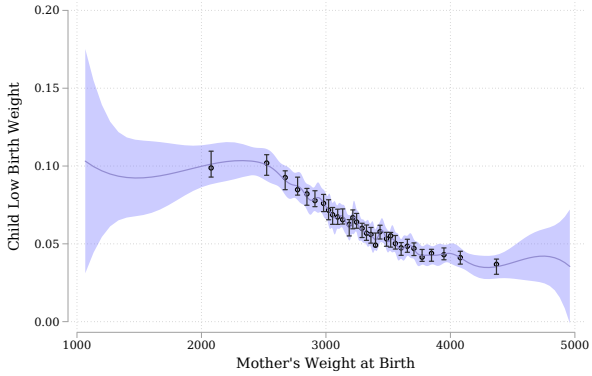
Many of these policies are additionally targeted at children who are born at less than 32 or 33 weeks of gestation, suggesting that the discontinuity may only bind for babies born at 32 weeks or above. Indeed, this is an argument made by Bharadwaj et al. (2013), who note that often policy documents or formal selection

[6] Searches in this document suggest 24 occurrences of the use of a 1,500 grams, while other arbitrary cut-points are not similarly prevalent (eg 4 mentions of 2,000 grams and only 1 mention of 2,500 grams which is the cut-off for definitions of low birth weight (LBW).

Figure 1: Intergenerational Transmission of Early Life Health Measures



(a) Maternal birth weight and children's birth weight



(b) Maternal birth weight and child's low birth weight status

Notes: Each sub-plot documents average outcomes of individuals born to mothers whose birth weight is indicated on the horizontal axis. Panel (a) considers the birth weight of second-generation children, while panel (b) considers the average proportion of second-generation children who are low birth weight (weight $<$ 2500 grams). Optimally spaced bins and their confidence intervals are documented as black points and error bands. A cubic B-spline and its 99% confidence interval is overlaid as a continual solid line and shaded area. Optimal definitions and recommendations follow Cattaneo et al. (2019a).

criteria for programs also note extreme prematurity along with very low birth weight. While this is often the case, it appears that it is not uniformly so. For example, in recent technical guidelines, a number of treatments are indicated as owing exclusively to the 1,500 gram threshold, and not based on gestational limits (for example, temperature regulation procedures discussed in Mulhaüsen Muñoz and González (2016)). As we discuss later in the paper, we also observe evidence suggesting elevated rates of hospitalization for babies born just below 1,500 grams even among individuals born at 32 weeks or earlier. Thus, while it is clear that treatment rules will bind most cleanly for babies born below 1,500 grams but at 32 weeks of gestation or greater, it is not necessarily clear that no discontinuity will exist around the 1,500 gram cut-off for babies born at below 32 weeks. Thus, as we discuss in section 4.2, while we generally follow Bharadwaj et al. (2013) in focusing on births at above 32 weeks of gestation, we additionally show robustness to not conducting this sample restriction.

In considering intergenerational links in health at birth, a stylised fact is that there is considerable, but not complete, closure of gaps in health at birth across generations. In Figure 1 we document, using microdata on all intergenerational links between births in Chile occurring after 1992, that while there is a clear tendency of lower birth weight children to go on to have births which themselves have poorer average health outcomes at birth, the intergenerational links are considerably tempered, with all binned averages located above 3,000 grams. These binned scatter plots suggest that on average, individuals weighing as little as 1,500 grams have

children with birth weights of over 3,000 grams, with this value rising steadily, reaching about 3,400 grams on average for individuals who weighted 4,000 grams at birth. Thus, a considerable reversion to the mean is observed, suggesting the existence of some intergenerational dependence in health measures at birth, but with considerably cushioned impacts. Similar patters are observed when considering low birth weight indicators, with children of LBW mothers being more likely to themselves be LBW, but nevertheless, 90% of individuals whose mothers were LBW "graduate out" of this status in the next generation.

3 Data

We generate matched microdata covering all the 6,617,638 births occurring in Chile between 1992 and 2018. These births are matched to their future survival history, inpatient hospitalization record, and any of their own births occurring in the future. In the case of any future births, we additionally observe the birth outcomes, and survival history of their children. For each birth we observe both child- and parent-level measures. Thus, these data are longitudinal covering up to 3 generations: characteristics of the mother and father of children born in generation 1, as well as characteristics of any births which occur later to children of women from generation 1. These generation 2 births allow us to observe the characteristics of babies from generation 1 when they go on to have their own children. In total, of the 6,617,638 births occurring between 1992-2018, 3,240,874 (48.9%) are girls, 435,014 of whom go on to have their own future live birth.[7] For these births we thus observe their outcomes at birth and in early life, their mother's outcomes at birth, early life and at the date of their child birth, and their grandparents' characteristics at the moment of their child birth.

We follow Lu and Vogl (2022) in referring to an intergenerational link as a lineage. We thus observe lineages which consist of mothers and their children, where mothers are observed both when they were born, in which case information on their mother (the "grandmother") is observed, and if they give birth, are additionally observed at this time, in which case information on their children is observed. We lay out this nomenclature, and the structure of lineages in our data in Appendix Figure A1. Birth certificate records which are provided by the Ministry of Health contain each mother's national identification number, but do not contain the father's national identification number. For this reason, we are only able to work with links between a *mother's* health across generations as matching boys to their children is impossible. Available evidence suggests that father's health is also transmitted to future generations (Giuntella et al., 2019), which is a point we return to when considering total policy effects in Section 5.3.1.

[7] Given the period covered, these births are all occurring to women born between 1992 and 2006. Later-born individuals are still children for the full data coverage period. We discuss data matches and inter-generational links in the paragraph heading "Sample Matches" below.

Birth, Death and Hospitalization Data Birth registries in Chile are universal, estimated to cover 99.9% of all births. Birth registries contain high quality records of birth weight (measured in grams), gestational length (in weeks), and size at birth (in cms), as well as information on the place of birth, and mother's and father's education and employment, along with other covariates.[8] Individuals are recorded using their national identity number, assigned at birth. Data on these births are merged (using a masked version of the national identity number) with the hospitalization registry and the death registry, which cover all deaths and in-patient hospitalizations in the country. In total, 58,993 births are matched to a death record before the age of 1 year. And in total 83,841 births are observed to appear in the death registry at any point during this period.

For a number of mechanism tests, we consider hospitalization records. Prior to 2001, individual identifiers are missing from a considerable proportion of micro-level registries of hospitalizations (inpatient records). Given this, we do not consider hospitalization data prior to 2001, instead consistently working with subsets of birth cohorts for whom hospitalization records are complete. Thus, when we examine impacts of early-life health investments on later life health-care usage, we work with age-specific outcomes. For example, when considering hospitalization at age 1, we can examine this for cohorts whose hospitalizations are observed completely at this age, namely individuals born from 2001–2017. And when considering hospitalization at age 2, we work with the sample of births from 2000-2016, and so forth for other ages. Of all births, 2,924,796 are matched to at least one hospitalization, and in total 5,654,411 hospitalizations are matched with births, implying that the average number of hospitalizations per matched birth is 1.93. These hospitalizations cover all inpatient care provided in the country, both in public hospitals and private clinics, and include information on both the reason for hospitalization (recorded by standardized ICD-10 codes), as well as the duration of hospitalization in nights.

Additional Data In a number of cases, we consider parental and family responses to child birth. In these cases, we generate information on changes following observed births, namely whether families go on to have additional births and if so the birth spacing in months, whether children are observed to be covered by private, rather than public, insurance or are treated in private, rather than public, hospitals, and whether parents are observed to leave or join the labour market, or move to a higher wage industry.[9] In each case, these measures are generated from administrative registers, and so are subset to individuals which appear in the relevant registers. For example, in the case of private insurance and hospitalization, this measure is generated only

[8] These data do not, however, contain information related to factors such as BMI, weight gain during pregnancy, smoking, or other behavioural measures.

[9] Birth registries contain information on each individual's occupation. We cross each individual's occupation with information on the average salary and average hour worked in each occupation by region from large household surveys conducted every 2-3 years (Chile's CASEN survey), thus allowing us to measure average conditions in the industry in which mothers and fathers work.

for children who are observed to be hospitalized. Similarly, in the case of birth spacing and labour market changes, these measures are observed only for individuals whose parents go on to have a subsequent birth, as both of these measures are generated based off information contained in birth registries.

Table 1: Summary Statistics – All Births

	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A: First Generation Births					
Gestation Weeks	6,574,158	38.64	1.82	15.00	44.00
$\geq$ 32 Gestation Weeks	6,574,158	0.99	0.10	0.00	1.00
Birth Weight in Grams	6,576,104	3323.61	539.29	90.00	6480.00
Birth Weight < 1,500	6,586,950	0.01	0.10	0.00	1.00
Birth Length in cm	6,570,261	49.37	2.58	9.00	59.00
Death within 1st Year of Birth	6,328,288	0.00	0.01	0.00	1.00
Days spent in hospital by year 1	838,942	10.31	23.36	1.00	365.00
Number of admissions to hospital by year 1	4,221,975	0.28	0.72	0.00	81.00
Mother's Age	6,586,950	27.15	6.57	15.00	45.00
Mother's Education Years	6,572,906	11.43	3.25	0.00	21.00
Panel B: Second Generation Births					
Gestation Weeks	425,183	38.56	1.94	16.00	43.00
$\geq$ 32 Gestation Weeks	425,183	0.99	0.11	0.00	1.00
Birth Weight in Grams	425,162	3280.26	535.04	151.00	6375.00
Birth Weight < 1,500	425,583	0.01	0.10	0.00	1.00
Birth Length in cm	425,155	49.10	2.62	16.00	59.00
Death within 1st Year of Birth	366,731	0.00	0.01	0.00	1.00
Mother's Age	425,583	19.55	2.69	15.00	26.00
Mother's Education Years	425,130	11.18	2.17	0.00	20.00

Notes: Summary statistics are displayed for all births occurring in the first generation (all births from 1992-2018), as well as those births matched to prior births (second generation births). The full sample consists of all births occurring between 1992 and 2018 in Chile from administrative data maintained by the Ministry of Health. Certain observations do not have information for certain variables based on exposure definitions: for example death within 1 year of birth is only defined for individuals aged at least 1 year, and number of hospitalization days is only observed if individuals are hospitalized.

Sample Matches Sample matches are generated based on an individual's national identity number (the RUT) which is unique, ubiquitous, assigned at birth and used throughout life. In Appendix Table A1 we document the matches across registers. As noted above, not all births are matched with hospitalizations given that many individuals are never hospitalized, and not all hospitalizations are matched to births, as many hospitalizations occur to older individuals who were not born during the period of 1992–2018. Of particular interest is the cross within the birth registry, identifying the universe of all lineages where mothers were born from 1992 onwards. These 435,014 lineages by definition must only occur to mothers aged at most 26 years (mothers who were born in 1992, and had a birth in 2018). We document the full matrix of mother-child birth years in Appendix Table A2, which makes clear the composition of our intergenerational sample. While this

is a restricted sample, a feature of this sample is that it is similar in age range to that of East et al. (2023), who observe women up to the age of 28 when considering intergenerational links. This allows for some comparability of our results with theirs based upon women's age. In the case of Chile, in the period under study this is approximately the median age (refer to Appendix Figure A2).[10]

Summary Statistics Summary statistics for the matched microdata of the first and second generations are provided in Table 1. We observe average birth weights of around 3.3 kg among the full sample, and slightly lower when considering second generation births. These second generation births will necessarily be born to younger mothers (with an average age of 19.6 years, compared with 27.2 years in the full sample). A low proportion of births are VLBW (around 1% of the full sample), and on average most pregnancies are taken to full term (around 38.6 weeks in both the first and second generation sample). In Table A3 we document identical summary statistics, but condition only on individuals born close to the 1,500 gram treatment cut-off. In this case we observe a much greater proportion of VLBW infants (given the sample definition), very premature infants, and deaths within 1 year of birth. In general, mothers in this sample are slightly older, at 28 years in the full sample.

4 Model and Methods

4.1 A Conceptual Framework for Understanding Intergenerational Spillovers from Early-Life Interventions

Previous literature in this setting (Bharadwaj et al., 2013; Chyn et al., 2021) has proposed a first generation model consisting of initial endowments, medical treatments (which interact with arbitrary treatment assignment cut-offs), and subsequent post-hospital investments. In considering intergenerational links in this setting, we can use these models as a starting point, however need to add an explicit framework to take into account fertility choices.

As in Bharadwaj et al. (2013), consider the observed birth weight of individual i, denoted BW_i, which imperfectly proxies unobserved health at birth:

$$BW_i = H_i + \varepsilon_i.$$

[10] It is important to note that East et al. (2023)'s sample is based on a policy expansion occurring between 1975 and 1988, and includes all cohorts up to the ages of 28. Thus, theirs is balanced in ages, restricting the years observed. Our sample is different in that we observe data on all births between years between 1992 and 2017 (ie balanced in years), and cohorts are thus observed to reach different ages, with the *earliest* births reaching 26 by the end our data (not all births). With present data availability, this is the limit of what we can observe. We return to discuss effects across age groups in Section 5.3.3 of the paper.

Additionally, consider the inputs received by an infant at hospital, D_i, which depend, decreasingly, on health at birth, as well as discontinuous treatment assignment rules:

$$D_i = g(H_i) + \kappa \cdot \mathbb{1}[BW_i < c] + \nu_i,$$

specifically, here neonatal health treatments are shifted upwards by some amount κ when an individual is born with a birth weight below the cutoff c. In this setting, initial medical care D_i is correlated with unobserved health measures not captured by birth weight, implying that estimates of the impacts of early life health investments conditional on birth weight will be biased in standard models. This leads to the RDD:

$$y_i = f(BW_i - c) + \alpha \cdot \mathbb{1}(BW_i < c) + X_i\beta + \xi_i, \tag{1}$$

where $f(\cdot)$ captures local relationships between the running variable (birth weight) and outcomes of interest y_i, specifically allowing for split local-linear or higher order polynomials, X_i is a vector of covariates, and ξ_i an unobservable error term. This design allows for the impact of medical treatments to be isolated from unobserved health stocks in the neighbourhood of c, given that crossing the threshold assignment causes a discrete jump in treatment $\Delta D_i = \kappa$, provided standard RD assumptions related to continuity of unobservables surrounding the cutoff point are met.[11]

However, understanding what α captures depends upon how posterior events and investments interact with D_i. Bharadwaj et al. (2013); Chyn et al. (2021) discuss this in a single generation setting which we lay out briefly here. We then extend this into a multigenerational setting, where there are clear implications of selection into the second generation which may interact with initial treatment.

To understand the impacts of treatment receipt on a treated individual during their own life, consider the stylised setting where the outcome of interest is an individual's (own) health. In a simple framework, along the lines of Grossman (2000), suppose that these time-varying health measures denoted H_{it} depend on initial endowments (H_i), medical intervention at birth (D_i), as well as subsequent parental investment.[12] As Bharadwaj et al. (2013) note, parental investment can interact with initial medical treatments in such a way to reinforce or compensate initial treatment receipt. Following their notation, parental investment is denoted $I_t^{post}(H, D, \zeta)$, capturing accumulated investments up to t which may be a function of initial health, treatment

[11] We discuss these assumptions, as well as how we can probe their validity, in the following subsection 4.2. There we lay out full estimation procedures related to the RD design in equation 1.

[12] In Grossman's canonical health capital model, health at time $t+1$ depends upon health at time t as well as (recursively) health at birth. Here given we are interested in estimating impacts of early life interventions we focus principally on health at birth. However, as we show in an online Appendix, we can incorporate dynamic health flows in this model without greatly altering the implications of this model for understanding RDD estimates.

at birth, and subsequent shocks ζ. Thus, 'first generation' health is modelled as:

$$H_{it} = \phi_t H_i + \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, \zeta_i) + X_{it}\beta_t + \upsilon_{it},$$

where it is clear that individual health at time t, H_{it} may depend on initial treatment directly, but also in the way that this treatment D_i interacts with subsequent parental behaviour.

As in Bharadwaj et al. (2013), while the RD isolates an exogenous shift in treatment intensity, the coefficient of interest from 1 when considering H_{it} as the outcome of interest captures the full policy impact up until t:

$$\widehat{\alpha} = \psi_t \cdot \kappa + \varphi_t \cdot \Delta I_t^{post}(c),$$

consisting of both structural effects from treatment at birth, as well as differences in posterior parental investments surrounding the treatment cut-off.

However, when we consider how early life treatments can have intergenerational consequences, this requires explicitly taking into account individual selection into appearing in the second generation. Specifically, we must consider that a first generation woman's fertility decisions may depend on her own initial inputs at birth, as well as the way these interact with future outcomes. Naturally, an individual will only appear in the second generation if a first generation woman has given birth. Thus, consider each woman's fertility decisions $fert_{it}$ which is a binary variable taking 1 if the woman gives birth, and hence has a child in the second generation, and 0 if she does not give birth. We can write this decision in terms of a latent variable:

$$fert_{it}^{*} = \gamma_1 H_i + \gamma_2 D_i + \gamma_{3t} I_t^{post}(H_i, D_i, \zeta_i) + X_{it}\pi_t + \iota_{it}. \tag{2}$$

Individuals for whom $fert_{it}^{*} > 0$ will have $fert_{it} = 1$, otherwise, $fert_{it} = 0$. This is a simple model of fertility behaviour in which fertility at time t is a function of an individual's health at birth, treatment receipt at birth D_i, as well as total parental investment up until t, which in turn may interact with initial medical treatment.[13] In turn, we may posit that birth weight of individual i's child (the second generation birth, who will be indexed j) can be described in the following way:

$$BW_{ij} = H_{ij} + \phi_t H_i + \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, \zeta_i) + X_{it}\beta_t + u_{ij} \tag{3}$$

which is observed only if $fert_{it} = 1$. Note that this function now depends on heritable factors, and so the

[13] While this is a simple model it provides a frame-work for understanding policy mechanisms, as discussed below. We document in Appendix B that it can be easily extended to depend on individual health at time t, making notation slightly more cumbersome.

mother's early life health, health interventions at birth, as well as posterior investments in the mother all may explain her child's birth weight.

Our interest in this paper is to determine the impact of intensive treatment at birth in the first generation on health outcomes of the following generation. Thus, our dependent variable of interest measures human capital at birth. To consider what an RD estimate captures where the outcome is birth weight of the second generation, we can write the expected birth weight of the second generation *conditional* on the child appearing in the second generation in terms of the well-known Heckman (1974) selection equation. For simplicity, we assume joint normality of ι_{it} and u_{ij} with an arbitrary covariance term ρ. Then, combining equations 2 and 3 we have:

$$\begin{aligned} E(BW_{ij}|fert_{it}=1) &= H_{ij} + \phi_t H_i + \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, \zeta_i) + \\ & \quad X_{it}\beta_t + \rho\sigma_u \lambda[H_i, D_i, I^{post}(H_i, D_i, \zeta_i)] \end{aligned} \tag{4}$$

where selection is captured by the inverse Mills Ratio λ, the standard deviation of u_{ij}, and covariance term ρ.

This selection equation has clear implications for the RD estimate in the case where outcomes capture health at birth of the second generation. While the regression discontinuity design levies locally exogenous changes in treatment for first generation individuals, the total policy-relevant treatment effect on the second generation will consist of multiple posterior interactions with the policy. Specifically, the coefficient of interest when considering second-generation outcomes such as birth-weight can be written as:

$$\widehat{\alpha} = \psi_t \cdot \kappa + \varphi_t \cdot \Delta I_t^{post}(c) + \rho\sigma_u \Delta\lambda[c, \Delta I_t^{post}(c)]. \tag{5}$$

Therefore, the reduced form impact of a mother i of child j crossing the 1,500 gram threshold is decomposed intro three terms. First, $\psi_t \cdot \kappa$ denotes the direct intergenerational transmission of improved health at birth of the mother, on to her children (the 'structural effect' of the policy). Second, $\varphi_t \cdot \Delta I_t^{post}(c)$ denotes the spillover on subsequent generations of parental investments in mother i which are sensitive to the policy. This (grand-)parental investment channel will capture any changes in first generation parenting as a result of treatment-receipt which are then transmitted to the second generation. Finally, the third composite term captures selection into the second generation, or a fertility channel of the policy. This term, $\rho\sigma_u \Delta\lambda[c, \Delta I_t^{post}(c)]$ makes clear that this selection may operate in a number of ways. Given that the inverse Mills Ratio in equation 4 contains D_i as well as $I^{post}(D_i)$, selection may owe to direct policy impacts (e.g. healthier individuals

as a result of intervention at birth may be more likely to take pregnancy to term), and/or selection may owe to changes in parental investments (e.g. compensating behaviour by parents may increase the likelihood of fertility). In both cases, we are concerned only with the way which treatment receipt changes this selection, as if fertility is unchanged as a result of the policy $\Delta\lambda = 0$, and the third selection term will be null. Note that, importantly, if fertility selection is relevant, the scale term ρ capturing unobserved correlations between birth weight and fertility likelihood will impact both the sign and magnitude of the selection term, underscoring the importance of understanding selection into fertility.

Thus, in what remains of the paper, the estimated coefficient $\widehat{\alpha}$ will be interpreted as the total policy relevant treatment effect, based on individual and parental response up until the moment that the outcome of interest is measured. When discussing policy impacts, we aim to document the importance of parental investment, and (when relevant) fertility selection channels, considering each of the elements laid out in 5. We discuss available measures to capture each of these channels in Section 5.2.

4.2 Empirical Strategy

Estimation and Inference Procedures We estimate equation 1, with the baseline specification following quite standard procedures. Namely, our principal model consists of estimating local-linear models such that $f(BW_i < 1500)$ allows separate linear functions on either side of the 1,500 gram assignment threshold. Bandwidths for these local linear regressions are chosen optimally following MSE minimisation criteria devised by Calonico et al. (2020), and principal models present robust bias-corrected parameters and inference methods of Calonico et al. (2014). A triangular kernel is used to weight observations by distance from the cut-off. We follow Bharadwaj et al. (2013) in defining specific details of the principal specification, diverging from this previous specification only where clearly justified arguments exist, such as in the use of optimal bandwidth selection procedures (rather than a fixed bandwidth of 200 grams), as these procedures have been largely developed and popularised following the publication of their paper. Thus, along with a local linear specification as the preferred specification, we control for covariates indicated by Bharadwaj et al. (2013) which are maternal characteristics (education, age, marital status), type of birth service (midwife or doctor), birth region, and year of birth. We additionally include a heaping control at 50 gram intervals, a point we turn to discuss below. While we aim to follow the models by Bharadwaj et al. (2013) as baseline specifications, we also document the stability of estimates to a range of alternative modelling considerations.

Estimation Sample As we noted in section 2, and as laid out in Bharadwaj et al. (2013), there is clear evidence of a sharp cut-off in treatment rules at 1,500 grams when babies are born at 32 weeks or above. In

certain cases, babies born at below 32 weeks will access the same treatment regimes, whether they are born at above or below the 1,500 gram threshold. In main estimates we thus focus exclusively on the sample of individuals born at 32 weeks and above, and in which case the discontinuity certainly binds. This sample covers 56% of all births within 200 grams of the 1,500 gram threshold. However, there is some evidence to suggest that at least in a more limited way, there is a discontinuity in treatment around 1,500 grams for babies born at below 32 weeks (see *eg* Appendix Figure A3). Thus, as a robustness check, we additionally present results from the main text covering the full sample of individuals, regardless of their gestational length.

Probing Identifying Assumptions and Robustness to Alternative Models Identification in a regression discontinuity design relies on the well-known continuity assumption related to unobservable factors around the cut-point (Hahn et al., 2001). If all unobservable factors are balanced on either side of the discontinuity, provided a small enough bandwidth is used for estimation, and provided that the cut-off generates a discontinuous jump in treatment, RD models can be used to isolate the impact of treatment on the dependent variable(s) of interest. As an initial (partial) test of identifying assumptions, we conduct balance tests on a range of predetermined observable characteristics of mothers, fathers, and births, *before* any differential treatment can be applied. Any lack of balance in these tests can be viewed as concerning for identifying assumptions in the RDD. We note that these tests will refer to mother and father characteristics of the *first generation* of mothers, as second generation mothers have been exposed to the policy, and thus balance tests may capture policy impacts, rather than any actual mis-balance prior to treatment assignment.

A much discussed concern in RD models where birth weight is the running variable is related to heaping of the running variable (Almond et al., 2010; Bharadwaj et al., 2013; Barreca et al., 2011, 2016). Birth weight is recorded when infants are weighed immediately following birth by their attending doctor or midwife, and rounding in these measures may occur for a number of reasons, such as sensitivity of electronic scales, or particular interactions with the 1500 gram cut-off. As we document in Appendix Figure A4, while the birth weight distribution appears very normal in the full population (panel (a)), heaps are observed at intervals of 100, and less-so 50 gram bins (panel (b)). As noted by Bharadwaj et al. (2013), in this context, heaping is correlated with demographic characteristics given that certain hospitals are more propense to round. We thus follow a number of procedures to ensure that heaping cannot explain the documented results. Firstly, in baseline models we consistently control for heaping using separate fixed effects for observations at each 50 gram bin, additionally documenting the stability of results to specifications without such controls. Secondly, we estimate Donut RD models, following suggestions of Barreca et al. (2011, 2016) which removes observations with weights very close to the cut-off point, gradually varying the width of the donut hole. Thirdly, we

consistently follow Bharadwaj et al. (2013)'s exact proposed main specification with mother and father level controls, which removes concerns related to demographic correlates of heaping (we also document approximate balance around the cut-off in our extended sample). Finally, we present results for 'placebo' estimates where a false discontinuity is assumed at alternative 100 gram bin intervals, to determine whether heaping could explain any systematically observed patterns when crossing 50 or 100 gram boundaries.

Where this heaping would be most concerning is it were indicative of manipulation of the running variable, with medical professionals being more likely to falsely assign particularly sensitive children to a very-low birth weight classification *in line* with perceived policy benefits. We note that empirically this does not appear to the case.[14] As discussed in section 2, the 1,500 gram cut-off is strictly enforced *below* 1,500 grams, implying that the heap at 1,500 grams does not receive treatment. Nevertheless, we formally test for manipulation into treatment following Almond et al. (2010) who suggests collapsing data at the gram level, and testing whether there are more observations just below the relevant cut-off (< 1500 grams) compared to those just above the relevant cut-off (> 1500 grams). All of these tests, providing support to our identifying assumption, are discussed in a section on robustness checks, 5.3.3, where we also document robustness to alternative polynomial orders given recent evidence of Pei et al. (2021), alternative RD procedures, and alternative kernel weighting schemes.

Principal Outcomes and Multiple Inference As we wish to consider how the impacts of early life health receipt accrue over life, and into future generations, we necessarily conduct multiple hypothesis tests considering multiple outcomes and a single treatment receipt. To avoid concerns that any findings will simply owe to inflated type I errors from repeated hypothesis testing, we proceed in two ways. Firstly, across all principal intergenerational measures we generate a single outcome index, following Anderson (2008). Secondly, within all classes of outcomes considered, we report both standard p-values, and q-sharpened p-values, which control for the false discovery rate.

5 Results

5.1 Intergenerational Impacts of Health Intervention at Birth

In Figure 2 we present visual RD plots, and in Table 2 we present formal tests for inter-generational transmission of intensive health treatments at birth. In each case, outcomes capture a child's health at birth,

[14] As well as the tests discussed in the text, In Appendix Figure A5 we provide plots of the average size of babies and the average gestational length surrounding the 1,500 gram threshold. We see very little evidence suggestive of noteworthy differences in these outcomes, which would be expected if medical professionals were systematically manipulating birth weights as a response to some other visual criteria.

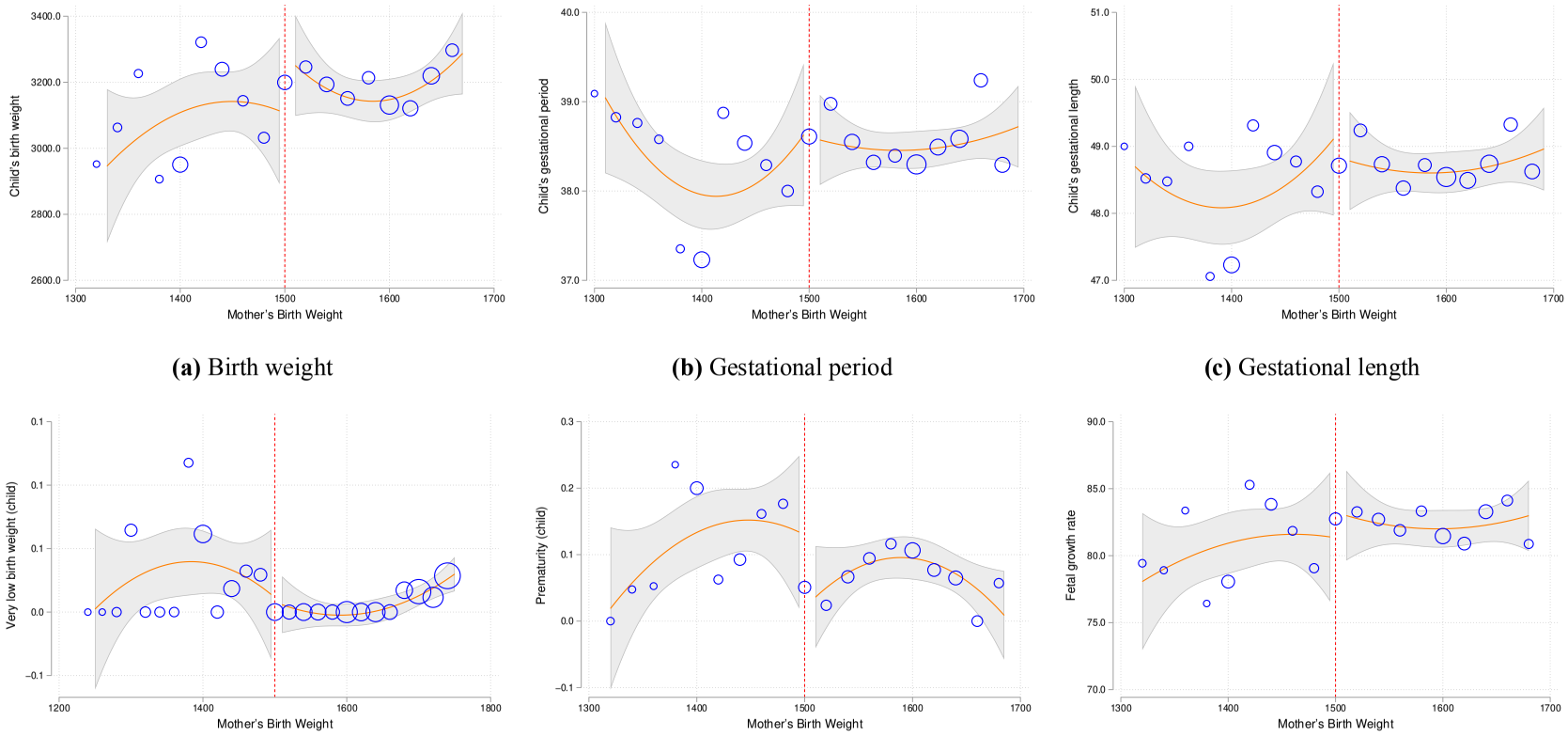
while the running variable refers to their *mother's* birth weight. Coefficients thus capture the impact of a mother having received intensive treatment when she was born on her children's health when she gives birth many years later. Figure 2 plots average outcomes in 20 gram birth weight bins within Calonico et al.'s Mean Square Error-optimal bandwidth around the discontinuity, overlaid on quadratic fits and their confidence intervals. Table 2 presents robust, bias-corrected estimates based on local linear regressions with a triangular kernel. The stability of these estimates to alternative functional forms, bandwidth and modelling choices is discussed in section 5.3.3 of the paper.

Across all outcomes considered, there is no evidence to suggest positive inter-generational gradients in treatment receipt at birth on children of the next generation. For example, in the case of birth weight, there is a clear *upward* shift when crossing from below the cut-off (more treated) to above the cut-off (less treated) individuals in Figure 2(a), and the point estimate in Table 2 suggests a return of -149 grams on the next generation. This point estimate is large, negative, though imprecisely estimated. While there are a large number of inter-generational links covered in these 30 years of data (421,382 in the case of birth weight for the 32+ week sample), when focusing on the optimal bandwidth of 171 grams either side of the cut-off, this is reduced to 811 effective observations.[15] Nevertheless, despite noisy estimates, in all measures observed in Panel A point estimates are consistent with there actually being negative transmission of early-life investments, rather than positive returns. As well as point estimates and standard errors (presented in parentheses in the table), we additionally present the p-value on a one sided test that corresponds to the program having negative inter-generational impacts. Thus, a low p-value would point to evidence to reject the null that the effect of the policy is negative, and high p-values can be considered as providing very little evidence to reject this null. Across all outcomes considered, these p-values range from 0.638 to 0.998, suggesting very little evidence to reject the null of a negative effect in favour of the alternative of a positive impact.

In considering well-being of infants at birth, it is important to consider measures beyond just birth weight (Conti et al., 2020). Point estimates in the case of gestational weeks point to an insignificant reduction by around 0.4 weeks, and in the case of birth length an insignificant reduction of around 0.2 cm. While most of these raw measures result in negative but insignificant inter-generational transfers, in the case of infant mortality of the second generation, we observe a marginally significant effect, consistent with individuals who receive intensive treatments going on to have children who are less likely to survive. This is documented in column 4 of Table 2, with the likelihood of death increasing by 0.007, or 0.7 per 1,000 births, which is approximately as large as the mean outcome in the inter-generational sample. The optimal bandwidth in this

[15] Power is a challenge in this setting. As we lay out in Appendix Figure A6, if the real effect were -100 grams, we would only have around 40% power to detect it. Given the structure of our data, 80% power is only achieved against a null of no effect if the true effect is ± 200 grams.

Figure 2: Descriptive Plots of Parental Policy Receipt and Child Health Measures



(a) Birth weight **(b)** Gestational period **(c)** Gestational length

(d) Very low birth weight **(e)** Prematurity **(f)** Fetal growth rate

Notes: Plots show separate quadratic fits estimated on each side of the 1,500 gram cut-off, in each case restricting attention to observations within the optimal bandwidth following Calonico et al. (2014). 95% confidence intervals of the quadratic fit are estimated, and circles represent average outcomes in 20 gram bins. The size of each point reflects the relative number of observations in each bin. Observations at 1,500 grams are plotted in binned averages, but are not used in estimating the quadratic fit.

Table 2: Intensive Health Investments and Birth Outcomes of the Second Generation

Panel A: Baseline Variables	Gestation (weeks) (1)	Birth weight (grams) (2)	Birth length (cms) (3)	Infant Mortality (4)
Birth weight < 1,500	-0.441	-148.907	-0.199	0.007*
	(0.502)	(96.771)	(0.563)	(0.004)
	[0.810]	[0.938]	[0.638]	[0.953]
Mean of Dep. Var.	38.422	3157.178	48.585	0.007
Observations	421,403	421,382	421,376	421,796
Optimal Bandwidth	196.6	171.2	193.5	142.3
Effective Observations	887	811	885	684
Observations (left)	305	284	305	251
Observations (right)	582	527	580	433
q-sharpened p-value	0.248	0.175	0.302	0.165
Panel B: Transformed Measures	Prematurity (5)	Very low birth weight (6)	Fetal growth rate (7)	Anderson Index (8)
Birth weight < 1,500	0.132*	0.045***	-2.736	-0.393*
	(0.072)	(0.016)	(2.010)	(0.220)
	[0.967]	[0.998]	[0.913]	[0.963]
Mean of Dep. Var.	0.088	0.015	81.830	-0.077
Observations	421,796	421,796	421,382	421,796
Optimal Bandwidth	189.8	257.7	178.2	232.2
Effective Observations	856	1,289	813	1,136
Observations (left)	294	361	286	347
Observations (right)	562	928	527	789
q-sharpened p-value	0.158	0.036	0.211	

Notes: Each column displays estimates of the change in the given dependent variable from above to below the 1,500 gram assignment threshold for mothers. In each case, local linear regression is used with a triangular kernel, calculating the MSE optimal bandwidth of Calonico et al. (2020). Robust bias corrected standard errors are reported in parentheses. Below standard errors, a one tailed t-test is calculated, which can be viewed as the support in favour of there actually being *positive* intergenerational transmission to the second genreation. q-sharpened p-values refer to corrections conducted across the entire class of outcomes. * p<0.10; ** p<0.05; *** p<0.01.

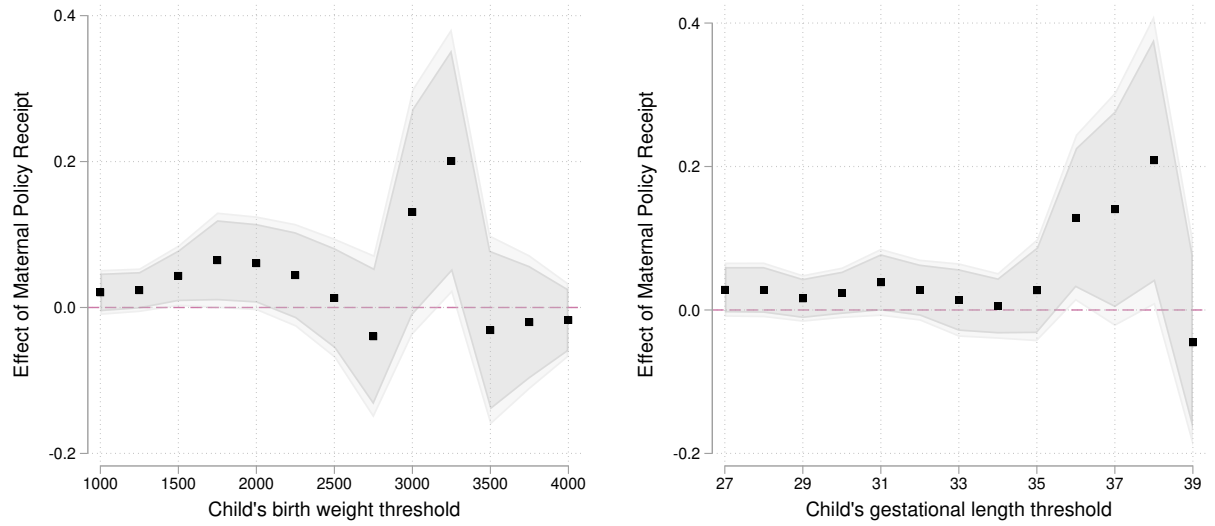
case is quite tight, at 142 grams surrounding the cut-off. This result suggests that the policy's reduction in infant mortality in the first generation (discussed in section 5.3.1 below), may actually be partially reversed in the second generation, though we note that this finding should be viewed with skepticism given that the FDR corrected p-value is only 0.165.

In Panel B we consider a number of alternative derived measures based on variables recorded in the birth register. In the case of prematurity and low birth weight, these are critical measures of health stocks at points quite low in the distribution of birth weight and gestational weeks. In this case, at low points in

the health distribution we find considerable negative returns to a mother's early life health receipt. In the case of both prematurity and low birth weight, we observe the individuals whose mothers just classified for treatment receipt are more than twice as likely to suffer from these conditions as individuals whose mothers were marginally above the treatment threshold. We similarly observe a reduction in fetal growth rate (defined as weight divided by gestational weeks), though this is not statistically significant. The final outcome in Panel B is an index of health measures at birth which is used to reduce the dimension of the test to a single dimension, and avoid inflated type I error rates. Here, in line with results observed throughout the table, this child health index is observed to be considerably worse among children of parents who received early life treatment, at around 0.4 of a standard deviation *lower* on average.

The results from Table 2 suggest that while tests on mean outcomes point to negative but imprecise results, particular key points of the distribution of these outcomes do see substantial shifts as a result of policy receipt. This is particularly the case in low birth weight measures, where we observe that a mother's receipt of intensive investments at birth makes her *more* likely to have low birth weight children. Given the high costs, both in terms of individual well-being as well as hospital-level costs for outcomes very low in the health distribution (Almond et al., 2010), we consider these distributional effects in a more flexible setting in Figure 3. In this case, each coefficient and confidence interval (90 and 95% CIs are plotted as shaded areas), refers to the probability that a mother has a birth with health measures *below* particular cut-offs indicated on the horizontal axis. Each coefficient and CI is thus generated from its own RD estimation, and is interpreted as the impact of policy receipt on intergenerational transfer of poor health indicators. The Figure 3a shows clear negative impacts of maternal policy receipt on her child's health. In proportional terms, this is particularly relevant in the far left tail of the distribution of birth weight. For example, in the case of 1,500 grams, we estimate that a mother who receives intensive early life health treatment increases her probability of having a birth at less that 1,500 grams by 4.5pp, which is an effective tripling considering that the proportion of such births occurring to mothers in this bandwidth of interest are around 0.015. Similar such patterns are observed when considering weights up to 2,250 grams. It is important to note here that this implies that (costly) early-life interventions in generation 1 imply a much greater proportion of births having to receive such interventions in generation 2. In the case of gestational length, Figure 3b, similar such distributional impacts are seen, with consistently positive probabilities of observing short gestational lengths, significant in the case of 32 weeks (extreme prematurity), with particularly large impacts observed at 37 weeks (prematurity). Again, these results are consistent with *negative* intergenerational gradients of health intervention at birth, and certainly never suggestive of positive gradients as frequently observed in prior literature.

Figure 3: Distributional Impacts of Early Life Health Interventions on Second Generation Health Stocks



(a) Birth weight

(b) Gestational length

Notes: Each point estimate (black square) and 90 and 95% confidence interval (dark and light shaded areas respectively) refer to RD estimates of the likelihood that a birth occurring to a treated girl born has health stocks (birth or gestational weight) below the cut-off indicated on the x-axis of each sub-plot. Each estimate is generated using the same sample and methods described in Notes to Table 2.

5.2 Understanding channels of inter-generational effects

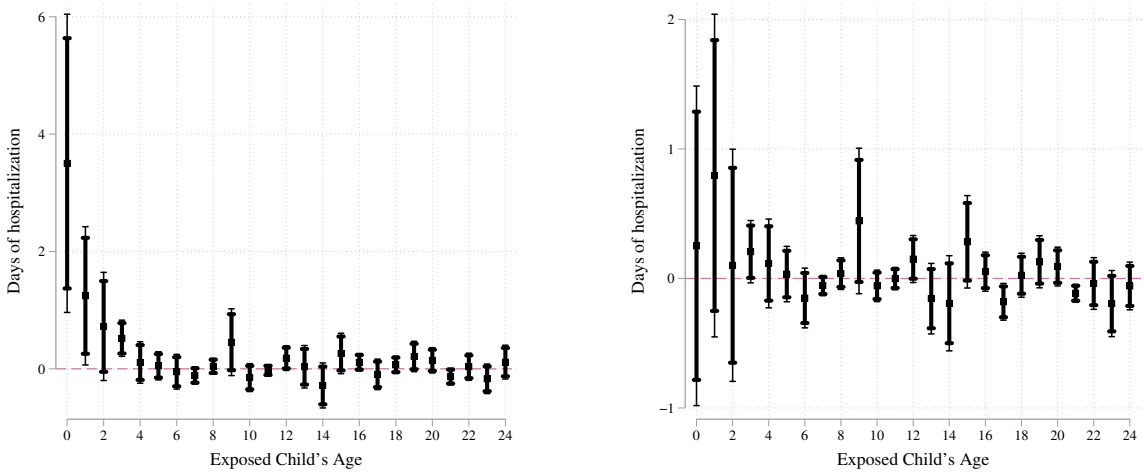
How can we rationalize the fact that we observe *negative* transmission of early life health interventions across generations when the broad consensus from extant literature is that these early life policies have broadly positive impacts on *first generation* individuals, and positive shocks are generally observed to transmit across generations? We turn to this question here, using the model laid out in Section 4.1 to guide potential competing explanations.

5.2.1 Stocks of Health in Real Time

One potential explanation for these results is that a mother's stock of health at birth is passed on to her baby (Appendix B, equation 10). Specifically, given the positive relationship between maternal health stocks and child health at birth (Lassi et al., 2013; Currie and Cole, 1993), for a health channel to explain *negative* intergenerational transmission, we would require that mothers who were marginally treated by the policy at birth be *less* healthy than mothers who were marginally untreated. While unlikely, such a phenomena could occur, if, for example, individuals who were less treated at birth were more likely to be hospitalized in early life, and ended up accruing more positive health stocks by maturity. Thus, the key consideration for this

mechanism is whether health stocks appear to be different, and in particular, whether there is evidence of compensatory investments in health during an individual's fertile period.

Figure 4: Long-Term Health Stocks and Early Life Interventions



(a) Hospital days

(b) Hospital days, 'chronic' causes

Notes: Each point estimate and confidence interval refer to the impacts of early life medical investment on an individual's days of hospitalization (panel (a)), and days of hospitalization related to chronic conditions (panel (b)). Thicker black error bars present 90% CIs, while thinner error bars report 95% CIs. Days of hospitalization (in both cases) are measured as totals for all individuals who have reached the age indicated, and take the value of 0 if the individual is not hospitalized this year, or otherwise a positive integer reporting the total number of days spent in hospital. All estimates follow the procedures laid out in section 4, and report RBC estimates using a local linear regression with a triangular kernel in the MSE optimal bandwidth.

In Figure 4 we examine the number of days an individual is hospitalized by year, estimated following the RD models laid out above at each age from 0 up until 24. Panel (a) of Figure 4 estimates the impact on the total number of days that affected individuals spend in the hospital for all causes, and panel (b) estimates the impact of early life medical treatment on hospital days which are classified as chronic based on ICD codes.[16] In both figures, it is apparent that there are relatively small or non-existent impacts on the long term usage of health care of having benefitted from early life health-care. Both when considering all hospital days, and hospital days for chronic causes, after the first few years of life there is little evidence of an enduring effect on hospitalizations.

We *do* observe a clear policy impact on hospitalizations early in life, estimating an increase of around 4 days in the year which birth occurs, 0.5-1.5 days per year up to year 3, and then smaller and non-significant impacts there-after. These effects are in line with those documented in Bharadwaj et al. (2013), reaffirming the power of the policy in terms of health investments. However, the lack of clear later life results suggests

[16]Chronic hospitalizations are classified using the Chronic Condition Indicator (CCI) developed by the Healthcare Cost and Utilization Project (HCUP).

that a role of direct transmission of health stocks from mother to children may be limited, given that we do not observe that individuals who marginally qualified for treatment have considerably different health stocks at maturity, at least if health stocks are proxied by hospitalizations.

5.2.2 Compensatory or Reinforcing Behaviour

A key channel which may explain null or negative intergenerational transmission of policy impacts relates to parental compensatory behaviour. Note in equation 5, that any change in parental behaviour as a *response* to treatment receipt at birth may act to counteract or reinforce any direct policy impacts. Specifically, a channel may exist in which parents whose children are born marginally above the 1,500 gram threshold and hence who are observed to be relatively worse off early in life, invest more heavily in these children, compensating initial disadvantage, and indeed fully closing the gap, explaining null or negative policy impacts in the long-run. This is something which can be empirically tested, if in place of examining the specific early life health measures in generation 2 births, we examine as outcomes in a RDD specification parental investment behaviours across the life of their children. The relevant test is then whether we observe evidence consistent with parents making more intensive investments to the right-hand side of the 1,500 gram cut-off, compensating the lack of policy receipt among these children.

To do this we collect a number of measures of parental investments in their children, or parental behaviours which may impact children's outcomes or well being. These are classified in terms of health investments, future demographic decisions, and in terms of parental labour market sorting decisions. Specifically, in the case of health investments, we examine whether, conditional on being hospitalized, children are covered by (more expensive) private insurance schemes, or are treated in private hospitals, which implies higher average out of pocket spending (Crispi et al., 2020). In terms of demographic decisions, we examine whether treatment receipt impacts future parental fertility behaviour, either making them more or less likely to decide to have future births, or changing the birth spacing in case of future births. In terms of labour market sorting, we test to see if *following on from their child birth*, mothers or fathers are observed to join or leave the labour market, to switch to a higher paying employment sector, or to switch to a less hour intensive employment sector. Such labour market responses could potentially impact child well-being on various margins: first an income margin if one or both parents opts to join to the labour market to afford greater investments in children, or second, a time investment if parents are observed to shift into less 'greedy' careers (as defined in Goldin (2021)), in favour of increasing (time) investments in children. Evidence of such labour movements and desires to seek careers which support both labour market and family investments are discussed, for example, in Goldin and Katz (2016).

Table 3: Parental Responses to Treatment Receipt

Panel A: Labour Market Responses

	Mothers				Fathers			
	Leaves Market	Joins Market	Expected Hour change	Expected Salary change	Leaves Market	Joins Market	Expected Hour change	Expected Salary change
Birth weight < 1,500	0.042	0.011	-0.275	77886.948**	-0.006	0.365***	-0.292	6798.468
	(0.056)	(0.038)	(0.292)	(31158.022)	(0.004)	(0.101)	(0.212)	(33673.318)
	[0.7742]	[0.6186]	[0.1727]	[0.9938]	[0.0885]	[0.9999]	[0.0841]	[0.5800]
Mean of Dep. Var.	0.230	0.237	40.528	5.4e+05	0.006	0.921	45.785	6.6e+05
Observations	366,044	669,779	558,787	554,751	856,752	47,842	1044260	1018479
Optimal Bandwidth	191.0	275.5	145.8	161.3	156.4	78.7	206.8	230.2
Effective Observations	834	2,097	877	1,001	1,438	38	2,396	2,703
Observations (left)	267	564	324	358	511	16	759	811
Observations (right)	567	1,533	553	643	927	22	1,637	1,892
q-sharpened p-value	0.549	0.725	0.549	0.046	0.362	0.003	0.362	0.725

Panel B: Parental Health Investments and Fertility Behaviour

	Private Hospitalization			Private Insurance			Future	Birth
	2	4	6	2	4	6	Birth	Spacing
Birth weight < 1,500	0.045	-0.042	-0.032	0.060	-0.165*	0.079	0.018	13.371
	(0.058)	(0.050)	(0.072)	(0.066)	(0.093)	(0.129)	(0.020)	(10.245)
	[0.7783]	[0.2007]	[0.3270]	[0.8167]	[0.0373]	[0.7302]	[0.8254]	[0.9041]
Mean of Dep. Var.	0.194	0.221	0.227	0.289	0.333	0.359	0.301	196.483
Observations	197,745	153,707	107,870	197,745	153,707	107,870	3556247	1363510
Optimal Bandwidth	151.8	142.4	176.5	186.1	108.9	146.7	234.7	168.5
Effective Observations	772	485	403	929	349	336	14,132	2,874
Observations (left)	277	178	131	317	134	118	4,113	1,005
Observations (right)	495	307	272	612	215	218	10,019	1,869
q-sharpened p-value	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes: Each column displays estimates of the change in the given dependent variable from above to below the 1,500 gram assignment threshold for newborns. In each case, local linear regression is used with a triangular kernel, calculating the MSE optimal bandwidth of Calonico et al. (2020). Robust bias corrected standard errors clustered at the gram level are reported in parentheses. p-values for one-sided tests are shown in square (H_1: negative) brackets. Significance stars for two-sided test: * p<0.10; ** p<0.05; *** p<0.01.

We lay out tests of parental responses to treatment receipt in Table 3 Panel A which tests for the existence of parental labour market responses and Table 3 Panel B which tests for the existence of changes in health investments or future fertility behaviour. In Panel A, RD estimates suggest relatively little evidence of affected mothers changing intensive margin labour supply decisions (leaving or joining conditional on their previous labour market status), however we do observe evidence consistent with mothers of treated children moving into higher wage sectors or participating in sectors which experienced higher wage growth. This is observed in column 4, with mothers of treated children being observed to change to sectors with mean monthly salaries which are approximately 80,000 CLP higher, around 100 USD based on current exchange rates. In the case of fathers, we observe relatively little evidence of similar changes in labour market circumstance. The one exception to this is an increase in the likelihood that fathers of marginally treated children join the labour market conditional on previously having been out of the labour market, though we note that this effect is

driven off a very small sample of fathers who previously had not participated in the labour market, and so is considerably underpowered. We note that in general, this result suggesting larger labour market responses for mothers than for fathers in the face of child health shocks is consistent with findings in the broader literature, for example (Eriksen et al., 2021) who find maternal shifts in labour market choices following negative child health shocks, and a broad literature, lead by Kleven et al. (2019), which documents relative inflexibility of father's labour supply to child birth, a phenomena also noted in Chile (Berniell et al., 2021). While these parental labour market responses to early life investments are of interest in their own right, for our results here, if anything they suggest that the observed patterns cannot be explained by labour market results, as mothers of individuals born just below the 1,500 gram threshold are observed to move to industries with higher, rather than lower salaries, without being much greedier in terms of time demands.

In the case of parental health investments and fertility responses (Table 3, Panel B), we observe relatively little evidence suggestive of consistent changes in the way which parents invest in their children's health care. This coheres with evidence presented by Bharadwaj et al. (2013) who found relatively little evidence of changes in educational investments by parents in marginally treated versus marginally untreated children. Early in life, we observe that conditional on hospitalization, rates of private hospitalization and private insurance coverage did not significantly differ across the 1,500 gram threshold. Similarly, we observe no evidence to suggest that parents of treated individuals were more likely to go on to have another birth, or change the timing of future births. As a matter of fact, birth timing between a low birth weight baby and future births appears to be quite similar to birth spacing in the population (Appendix Figure A7).

In general, these results suggest that (first generation) parental responses can explain relatively little of the observed negative inter-generational impacts to children of the second generation. Had we observed clear evidence of compensating investments, where individuals just to the right of the cut-off received additional investments or otherwise more positive home environments, this may have suggested that initial favourable medical treatments of treated individuals were overwhelmed by later favourable investments in untreated individuals. If anything, we observe that the reverse may be true, given the relatively better labour market trajectories for mothers of treated children.

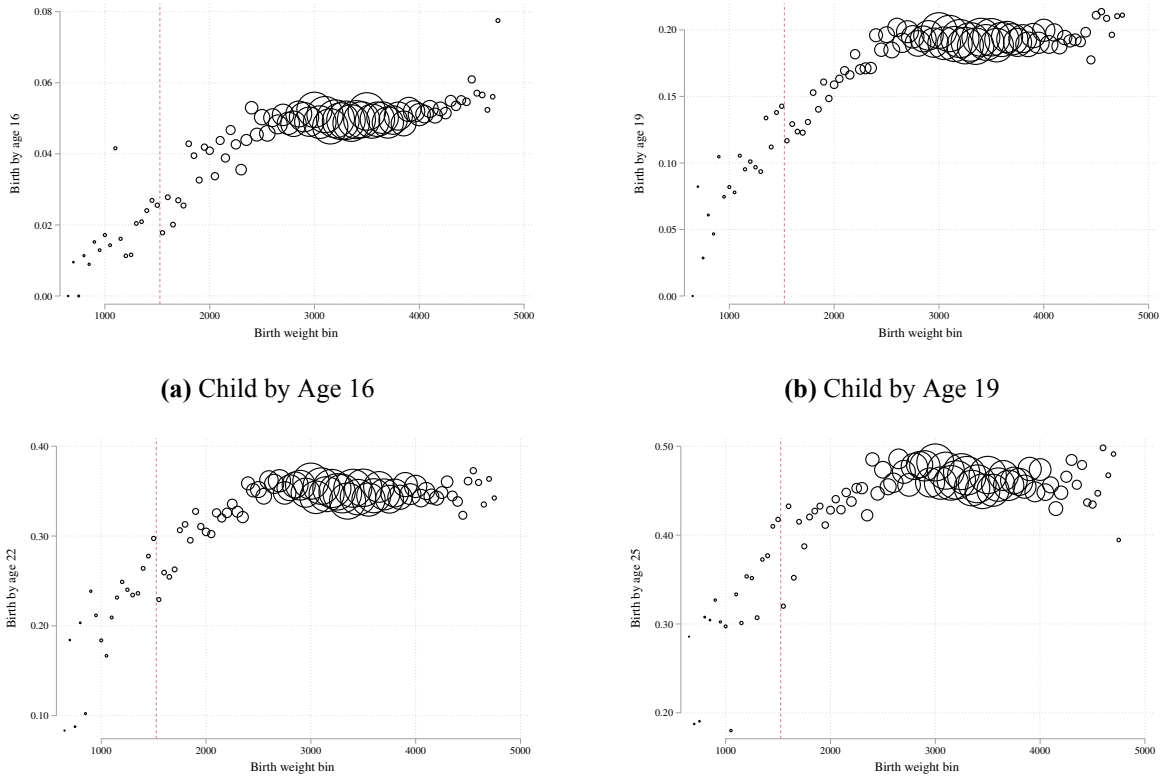
5.2.3 Selection

A key remaining consideration is selection into the sample. This is the third term in equation 5. This selection refers specifically to treatment-mediated *selection into fertility*, however we also note that selection could originate even at the time of birth of the first generation given that the treatment impacts survival of (potential) parents when they are born. Here we begin by focusing on fertility – a central point in our model

– before turning briefly to selective survival during infancy.

Selection into fertility Figure 5 presents descriptive plots of the likelihood that a woman gives birth by specific ages, graphing probabilities by an individual's own weight at birth. These plots are based on all women exposed to the possibility of giving birth – that is individuals who have reached the age under consideration in each plot, and hence potential 'second generation' mothers. These are thus all individuals who could potentially become mothers in our second generation sample, with plotted values representing the actual proportion of those women who actually do become mothers. Each point refers to the average in 50 gram bins, with the size of the point representing the number of individuals on which the average is based. Any 'selection' of interest for our estimated parameter — the third term indicated in equation 5 — would be apparent by changes in rates of birth just at the point where treatment targeting ends.

Figure 5: Fertility and Stocks of Health at Birth



(a) Child by Age 16 **(b)** Child by Age 19

(c) Child by Age 22 **(d)** Child by Age 25

Notes: Each figure plots the likelihood that a women born at a particular birth weight goes on to give birth by the age indicated in the plot caption. Points represent average proportions in 50 gram bins based on the women's weight when she was born. Averages are calculated from the full sample of women observed in the birth register, who have reached the age indicated in each plot. The size of each point refers to the relative number of women in this birth weight bin. The red dashed line indicates the 1,500 gram VLBW threshold used to assign discontinuous medical intervention at birth. Similar plots for all ages from 15–26 are provided in Appendix Figure A8.

Prior to considering selection *per se*, it is immediately notable and noteworthy that there is a steep gradient in the likelihood of giving birth by each of the ages documented in Figure 5 at lower points of the birth weight threshold. For each of the ages documented, there is a steep gradient in rates of birth up to around 2,500 grams. Very similar patterns are observed for all possible ages (provided in Appendix Figure A8, for ease of visualization, only 4 plots are displayed in the main text). For example, in the case of births by the age of 19, individuals who survive to age 19 and were born below 1,000 grams have less than a 10% chance of having given birth, rising to around 15% among individuals around 1,500-2,000 grams, before levelling off at 20% above around 2,500 grams. While points low in the birth weight distribution are based on few observations, the regularity of this pattern, both within and across age groups is clear, with this gradient consistently being observed, and consistently flattening out from above around 2,500 grams.

As far as we are aware, this stylised fact has not been previously documented in the economic literature, and the few medical studies which have documented correlations between birth weight and fertility are observational and based on relatively small convenience samples (Vikström et al., 2014; van der Pal et al., 2021; van Gendt et al., 2015), or indeed consider other birth outcomes such as preterm birth (Swamy et al., 2008). These results suggest considerable returns to birth weight in ways not previously considered in the economic literature. It also interacts with a broader literature on labour market returns to health at birth, and women's labour market returns in particular. If higher birth weight individuals have higher educational and labour market returns (Behrman and Rosenzweig, 2004; Bharadwaj et al., 2018), and at the same time birth weight is positively correlated with fertility, then negative links between fertility and labour market outcomes (Adda et al., 2017; Bloom et al., 2009) may partially obscure the full human capital returns to birth weight.

In Figure 5 it is additionally notable to the plain eye that there is an important discontinuity in rates of birth which are observed around the 1,500 gram cut-off. Even in using quite crude 50 gram bins in the spirit of capturing descriptive patterns, in each case clear sharp declines are observed in rates of birth when moving from just below 1,500 grams (marginally treated individuals) to just above 1,500 grams (marginally untreated individuals). For example, in Figure 5c, around 30 percent of individuals are observed to have a birth before the age of 22 in the 50 gram bin just below 1,500 grams, with this proportion falling to around 23 percent, or declining by approximately one quarter, when moving to the bin just above 1,500 grams. These values are identified formally using identical RDD methods as used throughout the paper in Figure 6a. Here, age-specific estimates are presented, estimated following equation 1, where the outcome considered is the total number of births a (potential) second generation mother has by age x.[17] We observe large, and generally

[17]This is a real integer, taking 0 if she has not had births, and the number of births if she has had 1+ births. Results are largely unchanged if instead of considering the total number of births, we consider a binary measure of having had any births by age x.

significant, impacts across all ages considered. By the age of 22, estimated impacts described in Figure 6a show that the number of births for individuals receiving the policy has increased by 0.1, against a base of approximately 0.25 births in individuals in the estimation bandwidth. Similarly large estimates are observed at each age above 20, growing to 0.2 additional births by the age of 25, versus a mean of around 0.35.

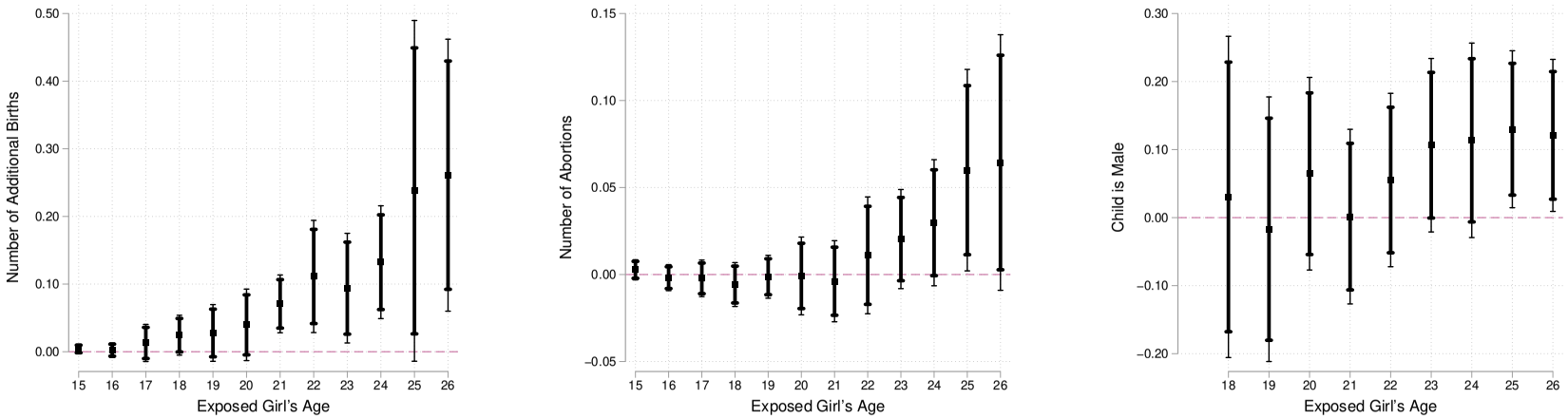
This result suggests that policy receipt has a clear long-run impact on treated girls. It makes them considerably more likely to give birth, as much as a quarter of a century after the initial policy receipt. This also provides a key explanation of the unexpected *negative* transmission of health at birth into the following generation. The policy may have a positive effect on fertility by rescuing marginal births of treated individuals. Individuals who in the absence of the policy would not have conceived or not have given live birth to a baby, do give birth to a baby when receiving intensive treatment at birth. This finding is consistent with the policy saving (second generation) babies which are marginally weaker, providing one explanation of the negative impacts observed in section 5.1 of this paper.

Fertility selection is thus a key mechanism which can explain the negative transmission across generations. We briefly consider what explains this fertility result (the 'mechanism of the mechanism'), prior to turning to potential alternative selection channels. In theory one could expect two main channels which could explain higher rates of childbirth following a shock such as policy receipt. Firstly, it may be the case that individuals are equally likely to conceive, but just less likely to take births to term when they do not receive the intensive early life medical investment. Or secondly, it may be that the early life medical intervention provides individuals greater social or reproductive resources to conceive. Due to the nature of administrative records, and high rates of miscarriage, we cannot observe all conceptions occurring around the cut-off, but rather only actual birth rates. As a proxy of miscarriage, or births not taken to term after conception, we can observe all spontaneous abortions which result in hospitalizations.[18] This is likely a considerable lower bound of all conceptions not taken term. We examine the impact of policy receipt on the likelihood of suffering a spontaneous abortion in Figure 6b. Analysis is conducted identically to the analysis of fertility. Here it is appears that based on *this* proxy, higher rates of birth among treated individuals do not owe to a greater likelihood of taking births to term, but rather, more likely due to a greater propensity to conceive.

However, an alternative and potentially more sensitive proxy of selective survival in utero is the sex ratio of births. It is well known that male fetuses are less likely to survive to birth, and are more sensitive to shocks in utero (see, eg Valente, 2015). Thus, an improvement in rates of survival of males at birth is evidence of improvements in rates of survival among relatively weaker babies (Trivers and Willard, 1973). We replicate

[18] These are observed in inpatient records, and are inferred from ICD-10 codes which are recorded as the reason for treatment. A full tabular classification is provided in an earlier version of this working paper (Clarke et al., 2022).

Figure 6: Impacts of Early Life Health Interventions on Fertility and Spontaneous Abortions



(a) Fertility **(b)** Abortions **(c)** Male child

Notes: Each point estimate and confidence interval refer to the impacts of early life medical investment on the number of births an individual has had by each age (panel (a)), the number of abortions observed in hospitalization data (panel (b)), and the likelihood that an observed birth is a male (panel (c)). Estimates in panel (a) and (b) are defined for all second generation individuals surviving to each indicated age, and so each estimation sample consists of all individuals. Estimates in panel (c) are limited to individuals who have had a birth by the age indicated on the x-axis, and so are limited to births occurring at (at least) age 18, as otherwise samples are too small to reliably estimate RDD models. Thicker black error bars present 90% CIs, while thinner error bars report 95% CIs. All estimates follow the procedures laid out in section 4, and report RBC estimates using a local linear regression with a triangular kernel in the MSE optimal bandwidth.

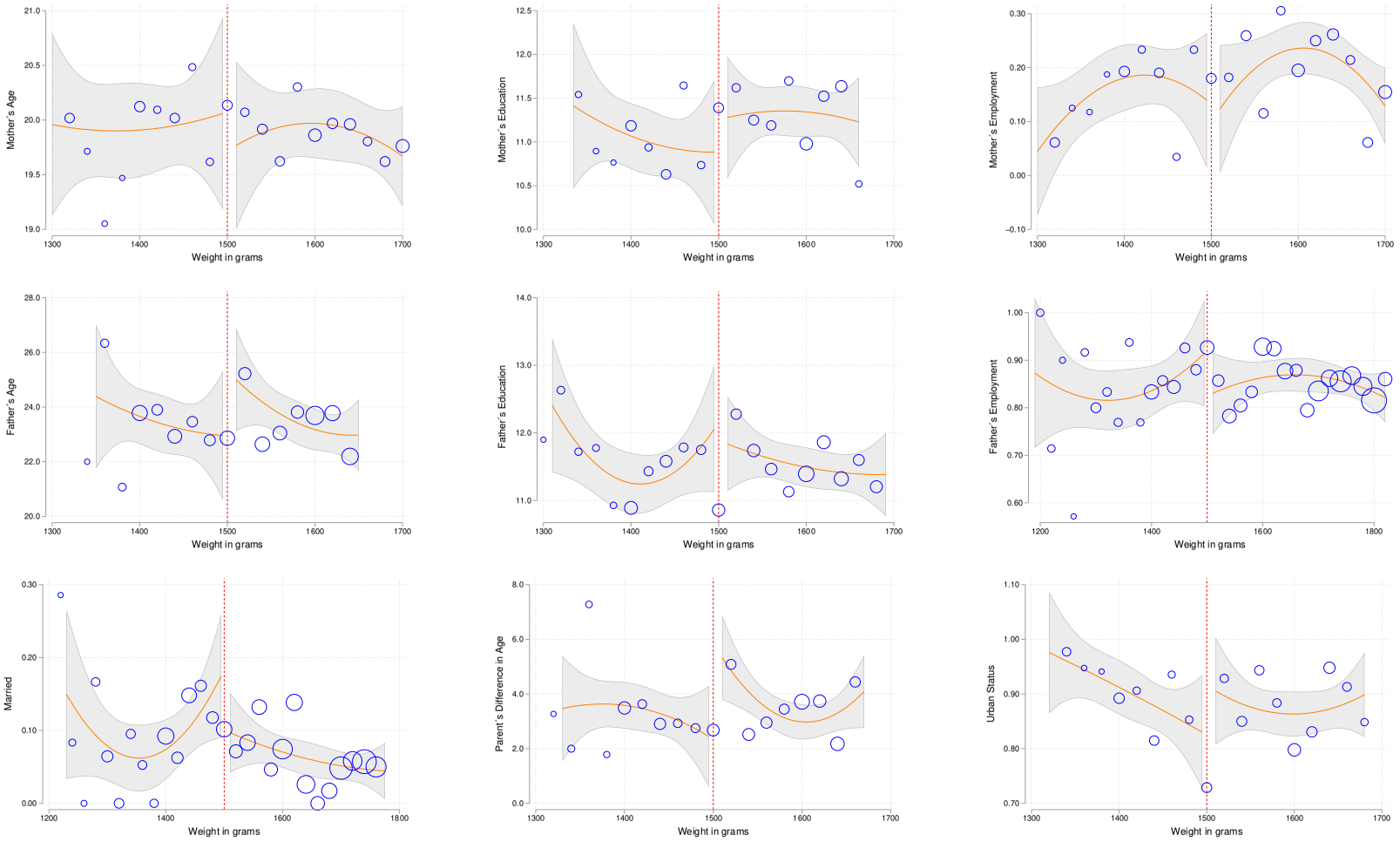
our analysis from panels (a) and (b) with the likelihood that a child is male in panel (c).[19] Based on this measure, we *do* see evidence in favour of fertility results being driven by selective survival of weaker fetuses in exposed mothers. In Figure 6c we observe substantial improvements in rates of birth of male children on the policy-exposed side of the RD threshold, and this becomes particularly noteworthy at ages where fertility increases most. These effects are large, suggesting around a 10 percent increase in likelihood of male children being observed among live births.[20] This suggests the observed results can, at least in part, be ascribed to birth selection occurring after conception, and potentially relatively early in gestation given a lack of clear impacts on spontaneous abortion.

In Figure 7, we examine whether this fertility selection is a generic result in the population, or something which is driven by particular groups. Formal RD estimates which correspond to each plot in Figure 7 are provided in Appendix Table A4. To test for selective fertility, we examine characteristics of mothers and fathers who go on to give birth in generation 2 around the 1,500 gram birth weight threshold. This figure is thus the intergenerational analogue of typical balance tests conducted in RDD to examine whether observations on each side of the cut-off are balanced, or are selected in someway. However, here, rather than acting as an identification check, these allow us to test for selective entry into the second generation of mothers. Across 9 outcomes observed in administrative data (mother's and father's education, age and employment, marital status, and whether births are multiple), we observe some evidence consistent with policy-driven fertility changes being more prevalent among certain groups. For example, we observe weak evidence to suggest that less educated women are slightly over-represented among treated rather than control individuals, which would be consistent with marginal births being more likely to survive among less educated women when they receive the treatment, compared with when they do not. We observe more clear evidence to suggest that treated individuals have *more* preferable partner matches when partners are observed: fathers are on average closer in ages to mothers, more likely to be employed, and mothers and fathers are more likely to married. Such a result may be consistent with treated individuals more generally forming better partnership matches as a result of treatment receipt. However, given that we do not observe partnership information for individuals who have not given birth, we cannot rule out that these results could simply imply that individuals with better partnership matches are more likely to have live birth when receiving treatment, compared to when not receiving treatment at birth.

[19] Note that because there are relatively few births before the age of 18 and bandwidths increase substantially, we focus only ages from 18-26.

[20] This is apparent event among descriptive statistics, with the likelihood of a birth being male in the 250 grams below the threshold being 0.47, compared with 0.51 in the 250 grams above the threshold.

Figure 7: Is There Policy-Driven Selection into Childbirth?



Notes: Each plot documents characteristics of mothers, fathers, or births of second generation births surrounding the 1,500 gram treatment application threshold of first generation births. A quadratic polynomial fit is graphed on either side of the cut-off, and points reflect average outcomes in 20 gram birth weight bins, with relative sizes reflecting the number of individuals in each birth weight bin. Each plot is displayed within Calonico et al. (2020)'s MSE optimal bandwidths. Formal tests of discontinuities at 1,500 grams are provided in Appendix Table A4.

Survival selection Beyond clear selection based on fertility, we consider the possibility that these negative inter-generational results simply owe to selection occurring due to survival at birth when intensive medical treatments were initially provided. As documented by Bharadwaj et al. (2013), and as discussed further in the following sub-section, this policy resulted in selective survival of individuals who received more intensive treatment, and as such, likely saved individuals with poorer health stocks. In order to consider whether this initial selective survival can explain our observed results, we conduct a counterfactual experiment where we impute non-surviving individuals on the right hand side of the cut-off, and consider whether specific life courses for these individuals could explain away the negative observed inter-generational transmission.

Specifically, the exercise consists of the following. Using estimates of the policy impact on infant mortality, we calculate the number of individuals on the right-hand side of the birth weight cut-off who we estimate would have survived had they received the treatment. This is estimated precisely using the RDD models discussed in this paper. We than impute the proportion of these individuals who would have given birth by each age using 20 gram birth weight bin-specific actual averages observed of this variable. We then add these additional 'counterfactual' births to our second generation sample (note that given different optimal bandwidths, the number of imputed individuals will vary slightly by outcome). As we do not know what their health outcomes would have been at birth, we consider a range of scenarios, imputing health outcomes across the actual distribution of health at birth.

The results of this exercise are provided in Table 4. This table makes clear that while selective survival of the first generation could affect the magnitude of the estimates, it would generally be insufficient to invert the sign of point estimates. For example, consider the case of child birth weight. If all counterfactual individuals on the right-hand side of the cut-off had a birth weight equal to that of the tenth percentile, the inter-generational effect could be as small as -89 grams. While if these non-surviving individuals had birth weight in the 90^{th} percentile, the true effect could be as large as -224 grams. This exercise is particularly clear in the case of a child's low birth weight status. Regardless of how extreme we make the counterfactual scenarios, the true effect remains large, and statistically significant, even at 1%. In Appendix C we consider a more complete selection exercise in which we vary both the health centile *and* the total fertility of imputed non-surviving individuals. These results similarly suggest that correcting for survival at birth in the first generation would not be sufficient to turn around observed results in all health dimensions.

5.2.4 Alternative Possible Explanations

There are a range of alternative possible explanations which we are not simple to test with the available data, but which can be illuminated by examining the available literature or aggregate statistics. This includes

Table 4: Selective survival and second generation outcomes – counterfactual analysis

		Counterfactual percentile assumed for survival				
	Baseline	10th	30th	50th	70th	90th
Panel A: Child's gestational length						
Birth weight < 1,500	-0.4411	-0.2591	-0.3822	-0.4930	-0.5160	-0.5805
	(0.5015)	(0.5112)	(0.5021)	(0.5003)	(0.4965)	(0.4986)
Effective Obs.	887	954	956	956	956	954
Imputed Obs.	0	79	79	79	79	79
Panel B: Child's birth weight						
Birth weight < 1,500	-148.91	-89.83	-135.37	-159.31	-175.70*	-224.30**
	(96.77)	(98.76)	(97.71)	(96.98)	(96.01)	(94.89)
Effective Obs.	811	876	845	845	845	845
Imputed Obs.	0	74	72	72	72	72
Panel C: Child's birth size						
Birth weight < 1,500	-0.1988	0.0653	-0.0712	-0.1821	-0.2929	-0.5140
	(0.5625)	(0.5712)	(0.5623)	(0.5619)	(0.5616)	(0.5615)
Effective Obs.	885	954	954	954	954	954
Imputed Obs.	0	79	79	79	79	79
Panel D: Child's infant mortality						
Birth weight < 1,500	0.007*	0.006	0.006	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Effective Obs.	684	745	745	745	745	745
Imputed Obs.	0	70	70	70	70	70
Panel E: Child premature						
Birth weight < 1,500	0.1323*	0.1350*	0.1350*	0.1350*	0.1350*	0.0012
	(0.0722)	(0.0722)	(0.0722)	(0.0722)	(0.0722)	(0.0736)
Effective Obs.	856	924	924	924	924	1,089
Imputed Obs.	0	77	77	77	77	93
Panel F: Child VLBW						
Birth weight < 1,500	0.0453***	0.0442***	0.0442***	0.0442***	0.0442***	0.0442***
	(0.0158)	(0.0158)	(0.0158)	(0.0158)	(0.0158)	(0.0158)
Effective Obs.	1,289	1,399	1,399	1,399	1,399	1,399
Imputed Obs.	0	122	122	122	122	122
Panel G: Fetal growth rate						
Birth weight < 1,500	-2.7356	-1.2524	-2.3924	-2.9518	-3.4530*	-4.7141**
	(2.0103)	(2.0418)	(2.0285)	(2.0164)	(2.0006)	(1.9804)
Effective Obs.	813	878	877	877	877	876
Imputed Obs.	0	74	74	74	74	74

Notes: Each panel displays outcomes under different counterfactual assumptions related to future outcomes for individuals who selectively did not survive birth as response to not receiving intensive treatment in generation 1. The left-hand panel replicates original estimates of policy receipt on inter-generational outcomes, and then additional columns impute outcomes for observations who selectively did not survive birth on the right-hand side of the treatment cut-off, assuming counterfactual outcomes at different percentiles of the health distribution at birth.

selective out-migration (a selection channel similar to those discussed above), educational attainment and school peer interactions, and other within-family spillovers.

One alternative explanation related to selection owes to migration out of the country. Such migration would need to be selective on one side of the treatment threshold. This explanation seems unlikely given that emmigration rates in Chile are very low, at 0.35 per 1,000 individuals, and generally concentrated at working ages, particular from 35 onward (Stefoni, 2011). A second explanation relates to individual interaction with treatment and the marriage market. One particular possibility is that individuals who are exposed to early life treatment have greater educational outcomes or contact with the educational system, and hence are more able to find partners. Impacts of early life medical receipt on education have been documented in a number of settings (Bharadwaj et al., 2013; Daysal et al., 2022), and so may partially explain observed results. While we cannot rule this out formally given that we do not observe partnership choices of individuals who do not give birth, we note that in Chile, secondary education is, at least by law, universal since 2003, suggesting that any such effects would need to be driven by partnership choices in higher education or among a more limited proportion of high-school deserters. Finally, there are precedents suggesting that early life health interventions at this margin can result in positive within-family spillovers, improving mother's mental health as well as sibling outcomes (Daysal et al., 2022). While our data does not allow us to consider such measures in this context, for this to explain a *negative* intergenerational health spillover, these positive family spillovers such as improved maternal mental health would need to map negatively into health of the following generation.

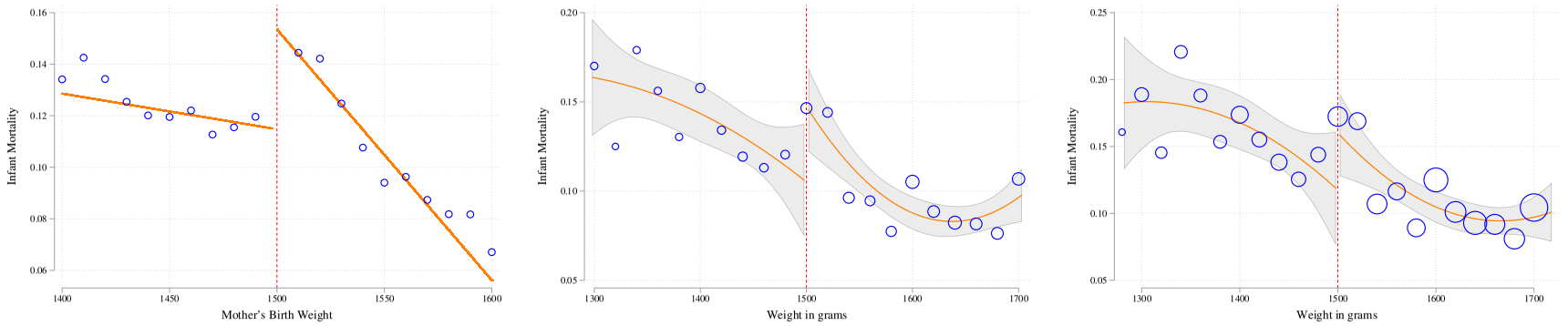
5.3 Supporting Assumptions and Robustness

5.3.1 Proof of Treatment

Prior to considering the robustness of main results to alternative specifications and modeling decisions, we seek to confirm previous evidence of the relevance of this treatment in this setting. This is a 'first-stage' which establishes the feasibility of later observing *inter-generational* spillovers of the initial treatment. Bharadwaj et al. (2013) have demonstrated that assignment to intensive neonatal care regimes due to crossing the VLBW threshold brought about sharp reductions in rates of infant mortality in Chile. We first document that these results can be replicated with the newly public matchable micro-date files, on various samples of data, and using both their original and more recent optimal RD methods. We also document that these results are found among earlier generations, which is key in demonstrating the relevance of these treatments among cohorts who will be mothers in the inter-generational sample.

Figure 8 presents regression discontinuity plots which examine binned rates of infant mortality among births occurring in Chile around the 1,500 gram cut-off. Panel (a) replicate Bharadwaj et al. (2013)'s methods and definitions, working with the same sample of births between 1992-2007. Note that here Bharadwaj et al. (2013) bin weights in blocks of 30 grams, centred on points of 10 grams, with the exception of the point clos-

Figure 8: Birth weight Assignment Thresholds and Infant Mortality



(a) BLN Method and Sample

(b) Full Period and Optimal Bandwidth

(c) Early Cohorts and Optimal Bandwidth

Notes: Each sub-plot examines the impact of crossing the VLBW threshold on infant mortality. All panels present estimates for gestational weeks 32 and above (where assignment rules clearly apply). Panel (a) replicates Bharadwaj et al. (2013)'s methods using overlapping (30 g) bins and a 100 gram bandwidth. Panel (b) uses optimal bandwidth selection methods of (Calonico et al., 2020), and plots a quadratic fit wth 95% CIs, where point sizes are indicative of relative sample sizes. Panel (c) replicates the optimal plot from panel (b) but focusing only on earlier birth cohorts, who are represented as mothers in the intergenerational sample.

est to the cut-off on either side, which is defined sharply. This implies that a single birth will be represented in multiple points. We follow their methods for comparability in panel (a), considering the same 100 gram range on either side of the cut-off, and in panels (b) and (c) use 10 mutually exclusive bins on either side of the treatment cut-off, and optimal bandwidth choices of Cattaneo et al. (2020). In all cases, a clear increase in infant mortality is observed when crossing the 1,500 gram threshold.

For our purposes, what is key is ensuring the relevance of initial treatment, which is clear even graphically in Figure 8. Formal RD estimates are presented in Appendix Table A5. Estimates are presented based on all three samples presented in Figure 8, estimating that, in the 'intergenerational sample', infant mortality falls by 2.5 pp, compared to a base of 11.2pp, or by approximately 22%, clearly establishing the relevance of these treatments in this particular setting.[21] If extending these tests to consider rates of infant death up to 2018, the results first documented in Bharadwaj et al. (2013) are, if anything, strengthened, with column 3 of Table A5 estimating a 2.7 pp reduction in rates of infant death against a base of 10.2 pp, even within the very tight bandwidth of 100 grams used by Bharadwaj et al. (2013), with very similar results observed in column 4 when estimating based on the optimal bandwidth of 134.4 grams. The relevance of the policy in providing initial treatment has also been established earlier in the paper when examining investments in hospitalizations, with sharp increases observed early in life (Figure 4). Note that these effect sizes are large, of the order of magnitude of a 20-30% decline in rates of infant death. A simple back of the envelope cost-benefit calculation presented in Online Appendix D suggests that these benefits from treatment receipt in the first generation dwarf any future costs incurred as a result of the policy given required investments in future generations. This holds even if we additionally allow for negative spillovers from fathers to their children.

5.3.2 Identification

Identifying the impacts of marginal medical investments at birth on subsequent outcomes based on our RD design requires that no other factors vary sharply local to the 1,500 gram cut-off. We consider a number of tests of these identifying assumptions, though note that this design has previously been validated in this and other settings by Bharadwaj et al. (2013); Chyn et al. (2021); Almond et al. (2010). In particular, our tests here focus on two considerations: firstly, is there balance at treatment receipt of observable characteristics, and secondly, is there evidence of manipulation of birth weights suggesting that there may be systematic, or strategic, selection into treatment by medical practitioners, or by families of babies born in this bandwidth.

A first consideration which is key is in examining the pre-determined characteristics of individuals who are located just below the 1,500 gram threshold, and hence in receipt of intensive medical intervention, and

[21] If we alternatively examine only earlier years, for example 1992–2001, similar results are observed. See Appendix Figure A9.

those located just above the threshold. We would be concerned if we observed that certain individuals are more likely to receive access to the policy, as it may illustrate imperfect compliance with the treatment threshold, and confound estimates presented throughout this paper. We examine a number of measures of first generation mothers, fathers and births, presented graphically in Appendix Figure A10, with associated RD tests in Appendix Table A6. Across outcomes tested, we observe relatively little evidence of mis-balance. Appendix Table A6 shows that across 10 outcomes, a single covariate is significant (at the 10% level only). Specifically, we do not see that treated individuals have parents of substantially different ages or educational levels or who are more closely aligned in age, nor do we see that they are more likely to have fathers recorded on the birth certificate, or are more likely to be born in urban areas. This balance point is also highlighted by Bharadwaj et al. (2013).

Nevertheless, a general concern in this case is still related to manipulation of the running variable, potentially not captured by these observable factors. If parents or medical practitioners were able to manipulate the running variable (officially recorded birth weight), they may be able to selectively ensure coverage for certain types of individuals. If this is the case, and if such manipulation owed to visual clues or information from patient histories inferred by medical practitioners, one may suspect that babies registered as having weights just to the left of the cut-off look different in birth size or gestational length than those just to the right. In Figure A5 we observe no evidence to suggest that this is the case. We also note that in the case of manipulation, one would expect a greater likelihood of observing births at points close below 1500 grams, and a lower likelihood of observing births just above 1500 grams, which is not something we observe. As proposed by Almond et al. (2011), we conduct tests formally examining whether there are more births observed in microdata registered as having birth weight just to the left compared to just to the right of the cut-off, and find no evidence to suggest that this is the case (Appendix Table A7).

Another specific concern related to manipulation and measurement in this setting, discussed extensively in the existing literature, is the presence of heaping in birth weight (Barreca et al., 2011, 2016; Almond et al., 2011). This can be observed when examining simple plots of the frequency of individuals observed at particular birth weights. In Appendix Figure A4 we observe that while birth weight is regularly distributed when zooming out across the entire distribution, regular peaks are observed at 50 and 100 gram bins, apparent when zooming in on birth weight in panel (b). Similarly, Bharadwaj et al. (2013) note that this rounding occurs differentially in certain types of hospitals, implying that it may be not be innocuous in estimation. We note that in the case of assignment rules, individuals will receive treatment only if they are observed to have a birth weight just *below* 1500 grams. Given this, throughout this paper we control flexibly for heaping and include controls for demographic factors proposed by Bharadwaj et al. (2013). But to ensure that our estimates are

not driven by heaping, specifically that at 1,500 grams, we estimate the 'Donut' RD models suggested by Barreca et al. (2011, 2016), removing observations which are located very close to the cut-off.

Table 5: Identification Considerations – Donut RD and Placebo Treatments

	Gestation Length (1)	Birth weight (2)	Size (3)	Infant Mortality (4)	Premature (5)	VLBW (6)	Fetal Growth (7)	Sex Ratio (8)
Panel A: Original Model								
Birth weight $<$ 1,500	-0.441	-148.907	-0.199	0.007*	0.132*	0.045***	-2.736	0.121**
	(0.502)	(96.771)	(0.563)	(0.004)	(0.072)	(0.016)	(2.010)	(0.057)
Panel B: Donut RDD								
Donut = 0	-0.441	-148.907	-0.199	0.007*	0.132*	0.045***	-2.736	0.121**
	(0.502)	(96.771)	(0.563)	(0.004)	(0.072)	(0.016)	(2.010)	(0.057)
Donut = 5	-0.458	-156.801	-0.243	0.005	0.126*	0.046***	-3.022	0.113*
	(0.543)	(103.584)	(0.614)	(0.004)	(0.075)	(0.016)	(2.093)	(0.066)
Donut = 10	-0.382	-156.671	-0.190	0.005	0.101	0.045**	-3.706*	0.103
	(0.535)	(106.656)	(0.622)	(0.004)	(0.070)	(0.018)	(2.186)	(0.066)
Donut = 20	0.078	-84.791	0.163	0.007	0.036	0.031**	-1.048	0.064
	(0.343)	(72.649)	(0.447)	(0.005)	(0.055)	(0.014)	(1.559)	(0.074)
Panel C: Alternative Cut-offs								
Birth weight $<$ 1,250	0.040	-11.011	-0.477	-0.061***	0.001	0.050	0.129	0.241*
	(0.422)	(156.808)	(0.577)	(0.015)	(0.056)	(0.031)	(3.695)	(0.124)
Birth weight $<$ 1,750	-0.060	1.297	-0.107	-0.005	-0.017	0.005	0.174	0.028
	(0.210)	(59.874)	(0.245)	(0.005)	(0.023)	(0.012)	(1.274)	(0.038)
Birth weight $<$ 2,000	0.105	32.322	0.144	-0.003	-0.002	0.002	0.524	0.040**
	(0.101)	(23.957)	(0.162)	(0.006)	(0.015)	(0.005)	(0.524)	(0.020)
Birth weight $<$ 2,250	0.042	-2.081	0.029	-0.003	0.005	0.003	-0.185	0.022
	(0.060)	(22.488)	(0.113)	(0.005)	(0.010)	(0.004)	(0.528)	(0.021)
Birth weight $<$ 2,500	-0.059	-38.707**	-0.091	0.002	0.010	0.003	-0.843**	-0.009
	(0.058)	(16.721)	(0.084)	(0.002)	(0.010)	(0.004)	(0.353)	(0.013)

Notes: Panel A replicates estimates from Table 2 for ease of comparison. Panel B estimates RDD models following 1, however removing observations within x grams of the cut-off, were x is indicated in the Donut hole value. Panel C estimates RDD estimates following equation 1, however here rather than using the true discontinuity at 1,500 grams, alternative discontinuities at other maternal birth weight points are considered. All additional notes follow those to Table 2.

These estimates are presented in Panel B of Table 5. Here we observe that when considering a Donut hole of 5 and 10 grams, findings are largely unchanged, and generally we observe that negative inter-generational transmission is maintained when considering health measures at birth, even when considering a larger donut hole of up to 20 grams. This is particularly clear in the case of very low birth weight thresholds. As an alternative model check, in Panel C of Table 5 we also consider whether similar discontinuities are observed when considering alternative birth weight cut-offs in place of the 1,500 gram threshold. Across 40 such placebo tests, we observe 5 test statistics which suggest the existence of a significant relationship at 10% or lower, which is in line with expected type I error rates. Here in general, the magnitudes of point estimates are also substantially lower.

5.3.3 Robustness

Throughout the main results section we have presented models based on optimal RD procedures following standard choices, such as the use of bias correction, local linear estimation procedures, and weighting with a triangular kernel centred at the cut-off point. To avoid concerns related to specification search, we have precisely followed Bharadwaj et al. (2013) in the use of control variables. Nonetheless, we document the robustness of principal results to alternative specifications or empirical decisions.

In Appendix Figures A11-A12 we document that results are broadly consistent across reasonable bandwidth choices, and evolve as expected as bandwidths grow considerably. We consider this in two ways: firstly (in Figure A11) we simply vary the bandwidth of data in which the RD model is fitted, using this same bandwidth for bias-correction. Generally speaking, across inter-generational outcomes, results are observed to be largest when focusing on a bandwidth more tightly bounded to the cut-off, and grow smaller only when the bandwidth is pushed up considerably, to above 300 grams. This value is well above optimal, and arguably local linear regression will begin to become more questionable as the bandwidth grows, and additional bias creeps into estimates. Secondly, in Figure A12 we document results where the data used to conduct bias correction is larger than the bandwidth, maintaining constant the ratio between bandwidth and bias correction as that which is obtained in optimal calculations from Calonico et al. (2014). Across specifications, we observe results which are consistent with larger results when using a smaller bandwidth, with results being attenuated as bandwidth grows. This is reassuring in that identification is local, and to the degree that wider bandwidths are used, more bias is expected.

More generally, we consider a range of alternative models varying, firstly the bandwidth selection and/or use of bias correction, and secondly the functional form of the running variable. These results are provided in Appendix Tables A8-A9. These consider a number of procedures for each inter-generational measure considered. First, we report robust-bias corrected results (replicating results from the main text). Second, we provide results using an alternative manner of selecting the optimal bandwidth. Recent results of Calonico et al. (2020) suggest that rather than MSE optimal bandwidths, one may wish to select bandwidths which minimize rates of error in hypothesis testing in RD procedures. We thus additionally present estimates based on these minimum coverage error (CE) bandwidths. Third, we present results using 'standard' RD models with conventional variance estimates, and fourth we present bias-corrected RD results, again using conventional variance estimates. Each of these results are presented for split linear and split quadratic polynomials to capture the relationship between outcomes and the running variable. Across models, results point to consistently negative intergenerational relationships (in the case of birth weight, prematurity, fetal growth rate, VLBW

status and the Anderson (2008) index), or negative but insignificant results in the case of gestational weeks and birth length. One outcome (infant mortality) suggests noisier results, with signs flipping in certain cases, suggesting that this outcome should be considered with some caution, given its rarity in second generation births.

Finally, it is worth noting a number of points related to the estimation sample. Firstly, as we noted in section 2, our estimation sample consists of all individuals born at weeks 32 and above, who are unambiguously exposed to the policy. We additionally present main results for the full sample of both individuals born at 32 weeks above, as well as individuals born below 32 weeks. In general, we expect results may be slightly attenuated, given that it is less clear that these individuals are exposed to all aspects of the policy. These results are provided in Table A10 and Figures A13-A14 for measures of second generation health at birth, and Figure A15 for impacts of policy exposure on fertility. Results are quantitatively similar, though slightly dampened in certain outcomes. Secondly, an important data constraint relates to the ages at which mothers are observed. One concern may be that this is simply a phenomenon related to teen birth, or younger births, and not representative of older births. While we cannot consider births outside the age group of our sample (maximum 26), and while results are necessarily noisier when stratifying by age, suggestive results presented in Figure A16 point to effects that are large both at younger and older ages among the groups considered, potentially limiting these concerns (at least within the age groups studied).

6 Conclusion

In this paper we document a long shadow to public policies, with the impacts of intensive medical care at birth found to have appreciable impacts as much as a quarter of a century later, and to be transmitted by recipients across generations. We document that unlike a large literature showing virtuous impacts of policies and positive shocks when passed from mothers to their children, we observe that children of treated mothers have worse measured birth outcomes such as low birth weight.

In examining policy mechanisms, we find that this unexpected result owes to selection into child-birth as a result of early life medical care. Girls who are born weighing just below 1,500 grams, and who receive intensive early life investments are much *more* likely to go on to have their own birth. We find that policy receipt in generation 1 makes individuals more able to give birth, but on average give birth to babies with weaker stocks of health at birth. We document a new stylized fact clearly linking birth weight to future fertility (both within and outside of the bandwidth considered in this study), and also make clear that this relationship is modifiable, as treatment receipt is observed to increase rates of birth by around 25% around VLBW cut-offs.

In many settings where concerns exist about low or below-replacement rate fertility, this suggests another social return to early life investments. Indeed both this finding and back of the envelope cost-benefit analysis suggests that while in this setting second generation health at birth is observed to be worse on average as a result of policy receipt, total policy benefits far exceed costs.

It is interesting to consider why these results are different to those generally found in the extant literature. One reason may owe to external validity. During the period under study Chile was a middle income country, and much of the literature examining intergenerational spillovers is drawn from a high income setting. However, there is reason to believe that these findings may be relevant in broader settings. Estimated impacts of this specific early life treatment on both first generation health and later life educational outcomes has been shown to be comparable in Chile to Norway, Denmark, and the US (Bharadwaj et al., 2013; Daysal et al., 2022; Almond et al., 2010), suggesting some external validity in other outcomes in this domain. The relevance of an extensive (fertility) and intensive (health at birth) margin policy response seems to be key in explaining differences with existing results. The only other study examining inter-generational spillovers of health policies of which we are aware (East et al., 2023) is focused on a specific low-income group of recipients and finds no fertility response, and subsequent positive inter-generational spillovers. The degree to which extensive and intensive margin policy effects come into play in broader settings seems to be a relevant open question.

These findings have implications related to the ways downstream health and welfare policies are defined following initial policy receipt. While a number of influential papers show that the impacts of early life treatments are unambiguously positive for their recipients (Bharadwaj et al., 2013; Almond et al., 2010; East et al., 2023), suggesting the need for compensatory policies for individuals who marginally miss out on such policies, our results also point to the importance of reinforcing investments, at least when considering the second generation of the original policy recipients.

References

J. Adda, C. Dustmann, and K. Stevens. The Career Costs of Children. *Journal of Political Economy*, 125(2): 293–337, 2017.

R. Akresh, S. Bhalotra, M. Leone, and U. Osili. First and second generation impacts of the biafran war. *Journal of Human Resources*, 2021.

D. Almond, J. J. Doyle Jr., A. E. Kowalski, and H. Williams. Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns. *Quarterly Journal of Economics*, 125(2):591–634, 2010.

D. Almond, J. Doyle, Joseph J., A. E. Kowalski, and H. Williams. The Role of Hospital Heterogeneity in Measuring Marginal Returns to Medical Care: A Reply to Barreca, Guldi, Lindo, and Waddell. *The Quarterly Journal of Economics*, 126(4):2125–2131, 10 2011.

D. Almond, J. Currie, and M. Herrmann. From infant to mother: Early disease environment and future maternal health. *Labour Economics*, 19(4):475–483, 2012.

S. Alvear, J. Canteros, J. Jara, and P. Rodríguez. Costos reales de tratamientos intensivos por paciente y día cama. *Revista Médica de Chile*, 141:202 – 208, 02 2013.

M. L. Anderson. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484):1481–1495, 2008.

A. Barr and C. R. Gibbs. Breaking the cycle? intergenerational effects of an antipoverty program in early childhood. *Journal of Political Economy*, 130(12):3253–3285, 2022.

A. I. Barreca, M. Guldi, J. M. Lindo, and G. R. Waddell. Saving Babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics*, 126(4):2117–2123, 10 2011.

A. I. Barreca, J. M. Lindo, and G. R. Waddell. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1):268–293, 2016.

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

I. Berniell, L. Berniell, D. d. l. Mata, M. Edo, and M. Marchionni. Gender gaps in labor informality: The motherhood effect. *Journal of Development Economics*, 150(C), 2021.

S. Bhalotra and S. Rawlings. Intergenerational persistence in health in developing countries: The penalty of gender inequality? *Journal of Public Economics*, 95(3):286–299, 2011.

S. Bhalotra and S. Rawlings. Gradients of the intergenerational transmission of health in developing countries. *Review of Economics and Statistics*, 95(2):660–672, 2013.

P. Bharadwaj, K. V. Löken, and C. Neilson. Early Life Health Interventions and Academic Achievement. *American Economic Review*, 103(5):1862–1891, 8 2013.

P. Bharadwaj, J. P. Eberhard, and C. A. Neilson. Health at birth, parental investments, and academic outcomes. *Journal of Labor Economics*, 36(2):349–394, 2018.

S. Black, N. Duzett, A. Lleras-Muney, N. Pope, and J. Price. Intergenerational correlations in longevity. Technical report, 2022.

D. Bloom, D. Canning, G. Fink, and J. Finlay. Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(1):79–101, 2009.

A. Bütikofer, K. V. Løken, and K. G. Salvanes. Infant Health Care and Long-Term Outcomes. *The Review of Economics and Statistics*, 101(2):341–354, 05 2019.

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

S. Calonico, M. D. Cattaneo, and M. H. Farrell. Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs. *Econometrics Journal*, 23(2):192–210, 2020.

M. D. Cattaneo, R. K. Crump, M. H. Farrell, and Y. Feng. On Binscatter. Staff Reports 881, Federal Reserve Bank of New York, Feb. 2019a.

M. D. Cattaneo, R. Titiunik, and G. Vazquez-Bare. Power calculations for regression-discontinuity designs. *The Stata Journal*, 19(1):210–245, 2019b.

M. D. Cattaneo, M. Jansson, and X. Ma. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531):1449–1455, 2020.

E. Chyn, S. Gold, and J. Hastings. The returns to early-life interventions for very low birth weight children. *Journal of Health Economics*, 75(C), 2021.

D. Clarke, N. L. Bustos, and K. Tapia-Schythe. Estimating Inter-Generational Returns to Medical Care: New Evidence from At-Risk Newborns. IZA Discussion Papers 15593, Sept. 2022.

J. Colmer and J. Voorheis. The grandkids aren't alright: the intergenerational effects of prenatal pollution exposure. CEP Discussion Papers dp1733, Centre for Economic Performance, LSE, Dec. 2020.

G. Conti, M. Hanson, H. Inskip, S. Crozier, C. Cooper, and K. Godfrey. Beyond Birthweight: The Origins of Human Capital. IZA Discussion Papers 13296, Institute of Labor Economics (IZA), May 2020.

D. Costa. Health shocks of the father and longevity of the children's children. Working Paper 29553, National Bureau of Economic Research, December 2021.

F. Crispi, A. Cherla, E. A. Vivaldi, and E. Mossialos. Rebuilding the broken health contract in Chile. *The Lancet*, 395(10233):1342, 2020.

F. Cunha, J. J. Heckman, and S. M. Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

J. Currie and N. Cole. Welfare and child health: The link between afdc participation and birth weight. *The American Economic Review*, 83(4):971–985, 1993.

J. Currie and E. Moretti. Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight. *Journal of Labor Economics*, 25(2):231–264, 2007.

J. Currie and M. Rossin-Slater. Early-life origins of life-cycle well-being: research and policy implications. *Journal of Policy Analysis & Management*, 34(1):208–242, 2015.

N. M. Daysal, M. Simonsen, M. Trandafir, and S. Breining. Spillover Effects of Early-Life Medical Interventions. *The Review of Economics and Statistics*, 104(1):1–16, 01 2022.

V. Duque, M. Rosales-Rueda, and F. Sanchez. How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data. Working Papers 2019-17, University of Sydney, School of Economics, Nov. 2019.

C. N. East, S. Miller, M. Page, and L. R. Wherry. Multigenerational impacts of childhood access to the safety net: Early life exposure to medicaid and the next generation's health. *American Economic Review*, 113(1): 98–135, January 2023.

T. L. M. Eriksen, A. Gaulke, N. Skipper, and J. Svensson. The impact of childhood health shocks on parental labor supply. *Journal of Health Economics*, 78:102486, 2021.

S. Eshaghnia, J. J. Heckman, R. Landersø, and R. Qureshi. Intergenerational transmission of family influence. Working Paper 30412, National Bureau of Economic Research, September 2022.

M. Feldstein. Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680, 1999.

J. L. García, J. J. Heckman, D. E. Leaf, and M. J. Prados. Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, 128(7):2502–2541, 2020.

O. Giuntella, G. La Mattina, and C. Quintana-Domeque. Intergenerational Transmission of Health at Birth from Mothers and Fathers. IZA Discussion Papers 12105, Institute of Labor Economics (IZA), Jan. 2019.

C. Goldin. *Career and Family: Women's Century-Long Journey toward Equity*. Princeton University Press, 2021.

C. Goldin and L. F. Katz. A most egalitarian profession: Pharmacy and the evolution of a family friendly

occupation. *Journal of Labor Economics*, 34(3):705–745, July 2016 2016.

R. Gonzalez, M. Merialdi, O. Lincetto, J. Lauer, C. Becerra, R. Castro, P. García, O. D. Saugstad, and J. Villar. Reduction in Neonatal Mortality in Chile Between 1990 and 2000. *Pediatrics*, 117(5):e949–e954, 2006.

M. Grossman. Chapter 7 – The Human Capital Model. In A. J. Culyer and J. P. Newhouse, editors, *Handbook of Health Economics*, volume 1 of *Handbook of Health Economics*, pages 347–408. Elsevier, 2000.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

J. J. Heckman. Shadow Prices, Market Wages, and Labor Supply. *Econometrica*, 42(4):679–694, July 1974.

M. E. Hübner, J. Nazer, and G. Juárez de León. Estrategias para Mejorar la Sobrevida del Prematuro Extremo. *Revista Chilena de pediatría*, 80:551 – 559, 12 2009.

J. Jiménez and M. I. Romero. Reducing Infant Mortality In Chile: Success In Two Phases. *Health Affairs*, 26(2):458–465, 2007.

H. Kleven, C. Landais, and J. E. Søgaard. Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11(4):181–209, October 2019.

M. Lahti-Pulkkinen, S. Bhattacharya, K. Räikkönen, C. Osmond, J. E. Norman, and R. M. Reynolds. Intergenerational Transmission of Birth Weight Across 3 Generations. *American Journal of Epidemiology*, 187 (6):1165–1173, 10 2017.

Z. S. Lassi, A. Majeed, S. Rashid, M. Y. Yakoob, and Z. A. Bhutta. The interconnections between maternal and newborn health – evidence and implications for policy. *The Journal of Maternal-Fetal & Neonatal Medicine*, 26(sup1):3–53, 2013.

C. Lee. Intergenerational health consequences of in utero exposure to maternal stress: Evidence from the 1980 Kwangju uprising. *Social Science & Medicine*, 119:284–291, 2014.

L. Lorenzoni and F. Koechlin. International comparisons of health prices and volumes: New findings. Technical report, OECD, 2017.

F. R. Lu and T. Vogl. Intergenerational persistence in child mortality. Working Paper 29810, National Bureau of Economic Research, March 2022.

C. Mardones and M. Riquelme. Estimation of the value of statistical life in chile and extrapolation to other latin american countries. *Latin American Research Review*, 53(4):815–830, 2018.

P. Mena Nanning, G. Mülhausen Muñoz, J. M. Novoa Pizarro, and G. Vivanco Giesen. Guías Nacionales de Neonatología. Technical report, Ministerio de Salud, 2005.

M. Mogstad and G. Torsvik. Family Background, Neighborhoods and Intergenerational Mobility. Draft chapter, handbook of family economics, 2022.

G. Mulhaüsen Muñoz and A. González. Guía de Práctica Clínica. Unidad de neonatología, Hospital San José, 2016.

C. B. Ogwulu, L. J. Jackson, A. E. Heazell, and T. E. Roberts. Exploring the intangible economic costs of stillbirth. *BMC Pregnancy and Childbirth*, 188(1):15, 2015.

M. A. Palomino, M. Morgues, and F. Martínez. Management of infants with chronic lung disease of prematurity in Chile. *Early Human Development*, 81(2):143–149, 2005.

Z. Pei, D. S. Lee, D. Card, and A. Weber. Local polynomial order in regression discontinuity designs. *Journal of Business & Economic Statistics*, 0(0):1–9, 2021.

P. Persson and M. Rossin-Slater. Family ruptures, stress, and the mental health of the next generation. *American Economic Review*, 108(4-5):1214–52, April 2018.

C. H. Rogers, F. J. Floyd, M. M. Seltzer, J. Greenberg, and J. Hong. Long-term effects of the death of a child on parents' adjustment in midlife. *Journal of Family Psychology*, 22(2):203–211, 2008.

M. Rossin-Slater and M. Wüst. What is the added value of preschool for poor children? long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics*, 12(3):255–86, July 2020.

H. Royer. Separated at Girth: US Twin Estimates of the Effects of Birth Weight. *American Economic Journal: Applied Economics*, 1(1):49–85, January 2009.

C. Stefoni. Perfil Migratorio de Chile. Technical report, International Organization for Migration, 2011.

G. K. Swamy, T. Ostbye, and R. Skjaerven. Association of preterm birth with long-term survival, reproduction, and next-generation preterm birth. *JAMA*, 299(12):1429–1436, Mar 2008.

R. L. Trivers and D. E. Willard. Natural selection of parental ability to vary the sex ratio of offspring. *Science*, 179(4068):90–92, 1973.

C. Valente. Civil conflict, gender-specific fetal loss, and selection: A new test of the Trivers–Willard hypothesis. *Journal of Health Economics*, 39:31–50, 2015. ISSN 0167-6296.

S. M. van der Pal, S. A. van der Meulen, S. M. Welters, L. A. Bakker, C. J. M. de Groot, A. H. van Kaam, and E. G. H. W. Verrips. Reproductive risks in 35-year-old adults born very preterm and/or with very low birth weight: an observational study. *European Journal of Pediatrics*, 180(4):1219–1228, Apr 2021.

A. W. van Gendt, S. M. van der Pal, W. Hermes, F. J. Walther, K. M. van der Pal-de Bruin, and C. J. M. de Groot. Reproductive outcomes of women and men born very preterm and/or with a very low birth weight in 1983: a longitudinal cohort study in the Netherlands. *European Journal of Pediatrics*, 174(6):819–825, Jun 2015.

A. S. Venkataramani. The intergenerational transmission of height: evidence from rural vietnam. *Health Economics*, 20(12):1448–1467, 2011.

J. Vikström, M. Hammar, A. Josefsson, M. Bladh, and G. Sydsjö. Birth characteristics in a clinical sample of women seeking infertility treatment: a case-control study. *BMJ Open*, 4(3):e004197, Mar 2014. doi: 10.1136/bmjopen-2013-004197.

M. Wüst. Universal early-life health policies in the nordic countries. *Journal of Economic Perspectives*, 36 (2):175–98, May 2022.

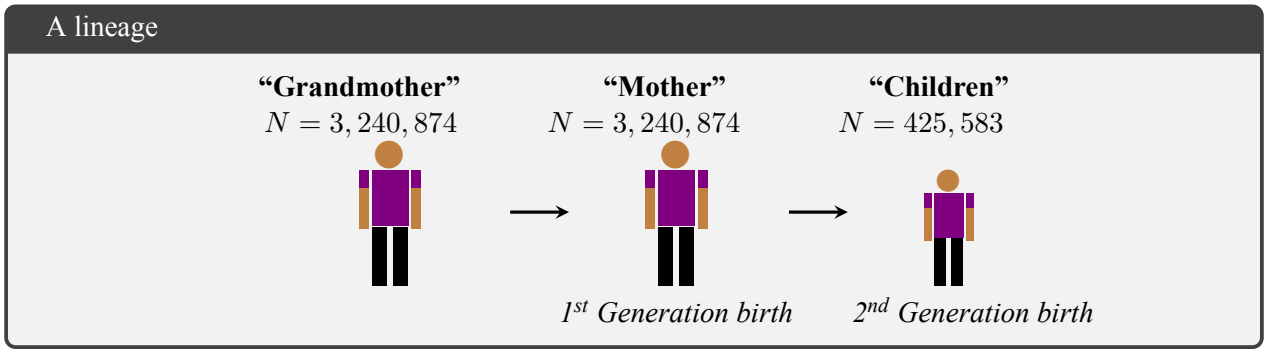
Online Appendices for:

"Estimating Intergenerational Returns to Medical Care: New Evidence from At-Risk Newborns"

Damian Clarke Nicolás Lillo Bustos Kathya Tapia-Schythe

A Appendix Figures and Tables

Figure A1: Schematic Representation of a Lineage



Notes: Birth certificates capture characteristics of the individual born, as well as their parents, and we observe the unique identity number of the parent as well as the individual born. Thus, for "Mothers", at their time of birth we observe their birth outcomes, as well as their mothers' characteristics ("Grandmothers"). While for "Children", we observe their birth outcomes, their mother's birth outcomes, and their mother's outcomes at the time the child was born. The nomenclature "mother" refers to all girls who could potential go on to being mothers, many of which have not yet given birth.

Table A1: Matched Observations between Microdata Registers

Register	Observations with Valid Unique IDs	Matched to Births	Matched to Hospitalization	Matched to Deaths
1992–2018 Births	6,617,638	435,014	5,654,411	83,841
2001–2019 Hospitalization	27,995,452	2,924,796	—	1,272,340
1992–2018 Deaths	2,393,583	83,841	4,627,999	—

Notes: Column 1 presents the total number of valid observations in each dataset. Column 2 notes the number of births which match to each dataset. In the case of the birth register, it refers to the number of births which match to other births in the data (ie mother–child links). Column 3 notes the number of hospitalizations which match to each other database. Note that in the case of births, the number of hospitalizations linked to births is not the same as the number of births linked to hospitalizations in the preceding column given that a single birth can be hospitalized multiple times. Finally, column 4 notes the total number of deaths which are matched with births occurring in the sample.

Table A2: Temporal Links between Mother–Child Matched Birth Years

Mother	Child															
Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1992	0	10	57	346	1,475	3,789	6,205	7,965	9,468	9,964	9,898	9,831	9,438	8,954	8,663	8,912
1993	0	0	6	60	399	1,622	3,806	5,869	7,632	8,962	9,344	9,826	9,413	8,578	8,162	8,260
1994	0	1	0	7	60	373	1,614	3,608	5,982	7,429	8,457	9,273	9,023	8,519	7,974	7,865
1995	0	0	0	2	11	76	418	1,437	3,498	5,560	6,654	8,000	8,298	7,770	7,480	7,235
1996	0	0	1	0	3	9	83	356	1,484	3,493	4,967	6,205	7,046	7,305	7,068	6,774
1997	0	0	0	0	0	0	10	61	400	1,442	3,183	4,684	5,433	5,687	6,123	6,478
1998	0	0	0	0	0	0	0	5	56	367	1,431	2,911	3,882	4,366	4,899	5,566
1999	0	0	0	0	0	1	0	2	7	52	369	1,331	2,349	2,914	3,415	4,081
2000	0	0	0	0	0	0	0	0	2	6	56	326	1,057	1,879	2,270	2,736
2001	0	0	0	0	0	0	0	0	0	0	3	49	261	754	1,297	1,638
2002	0	0	0	0	0	0	0	0	0	0	0	4	43	248	582	1,022
2003	0	0	0	0	0	0	0	0	0	0	0	1	7	23	166	471
2004	0	0	0	0	0	0	0	0	0	0	0	0	1	6	48	177
2005	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	28
2006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3

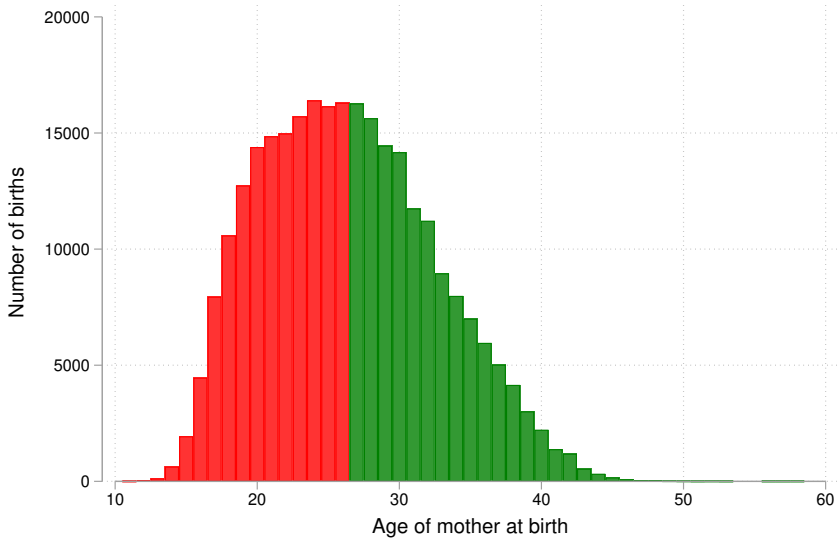
Notes: Values note the total number of mothers who are born in each of the years indicated in the left-hand column giving birth in each year indicated by row headers. All values are based on matched administrative records, and are provided from 1992 (the first birth year observed) up to 2006 (the last birth cohort in which a later-born child was linked to a mother born in this year.)

Table A3: Summary Statistics – Births local to the 1500 gram threshold

	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A: First Generation Births					
Gestation Weeks	26249	31.78	2.19	20.00	42.00
$\geq$ 32 Gestation Weeks	26249	0.53	0.50	0.00	1.00
Birth Weight in Grams	26263	1507.53	76.79	1367.00	1633.00
Birth Weight < 1,500	26263	0.44	0.50	0.00	1.00
Birth Length in cm	25859	40.32	2.09	19.00	56.00
Death within 1st Year of Birth	22734	0.00	0.02	0.00	1.00
Days spent in hospital by year 1	12436	39.59	37.27	1.00	365.00
Number of admissions to hospital by year 1	17532	1.26	1.43	0.00	43.00
Mother's Age	26263	28.02	7.13	15.00	45.00
Mother's Education Years	26077	11.45	3.37	0.00	21.00
Panel B: Second Generation Births					
Gestation Weeks	2909	38.35	2.11	22.00	42.00
$\geq$ 32 Gestation Weeks	2909	0.98	0.13	0.00	1.00
Birth Weight in Grams	2909	3176.81	557.94	363.00	4875.00
Birth Weight < 1,500	2913	0.01	0.12	0.00	1.00
Birth Length in cm	2909	48.67	2.80	20.00	56.00
Death within 1st Year of Birth	2471	0.00	0.00	0.00	0.00
Mother's Age	2913	19.92	2.67	15.00	26.00
Mother's Education Years	2909	11.16	2.21	0.00	18.00

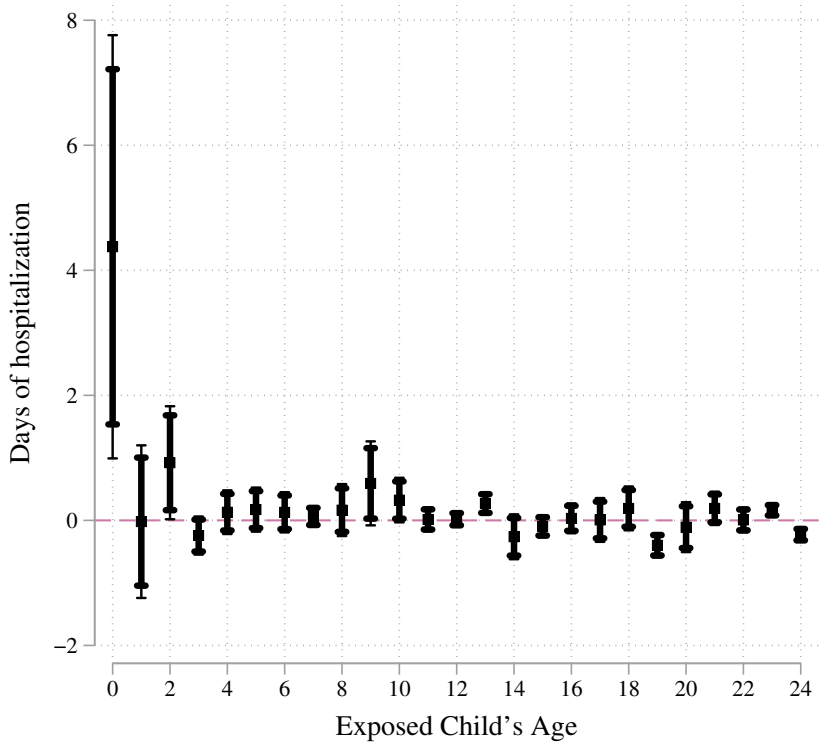
Notes: Summary statistics are displayed for births occurring close to the 1,500 gram treatment threshold for the first generation (all births between 1992 and 2018), as well as those births matched to prior births (second generation births). The full sample consists of all births occurring between 1992 and 2018 in Chile from administrative data maintained by the Ministry of Health, and here we subset based on treatment assignment (mother's birth weight). This is based on the 134.4 gram optimal bandwidth cut-off when considering infant mortality for the first generation. Hospitalization days refer to days *only* for those births which are admitted to hospital.

Figure A2: Age of Mothers at Childbirth



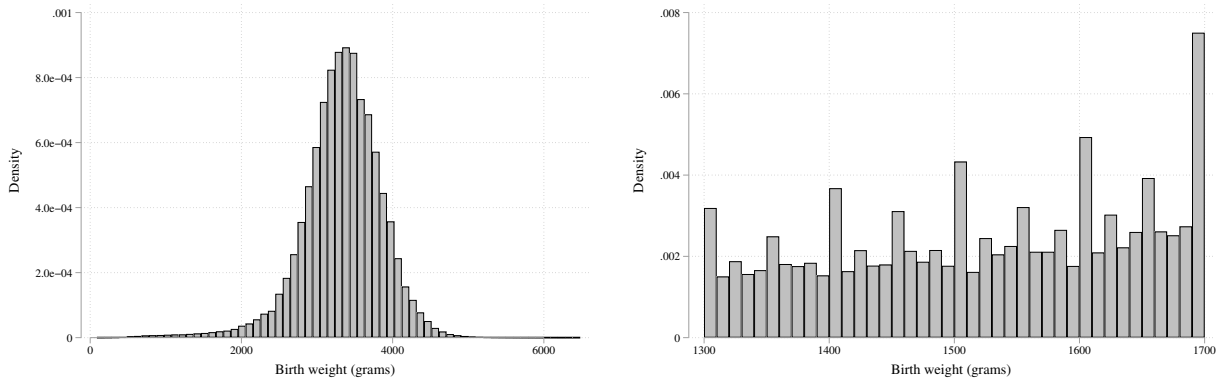
Notes: The distribution of births by mother's age at baseline (1992) is displayed. Births with mothers greater than age 26 are shown in green, while mothers aged 26 and below (corresponding to the maximum age in the intergenerational sample) are shown in red.

Figure A3: Discontinuities in Hospitalization in Babies Born at $<$ 32 Gestational Weeks



Notes: Each point estimate and confidence interval refers to the impacts of early life medical investment on an individual's days of hospitalization at each age, estimated by RD using the full sample of births *born at less than 32 gestational weeks* using the RD specification. Thicker black error bars present 90% CIs, while thinner error bars report 95% CIs. Days of hospitalization are measured as totals for all individuals who have reached the age indicated, and take the value of 0 if the individual is not hospitalized this year, or otherwise a positive integer reporting the total number of days spent in hospital. All estimates follow the procedures laid out in section 4, and report RBC estimates using a local linear regression with a triangular kernel in the MSE optimal bandwidth.

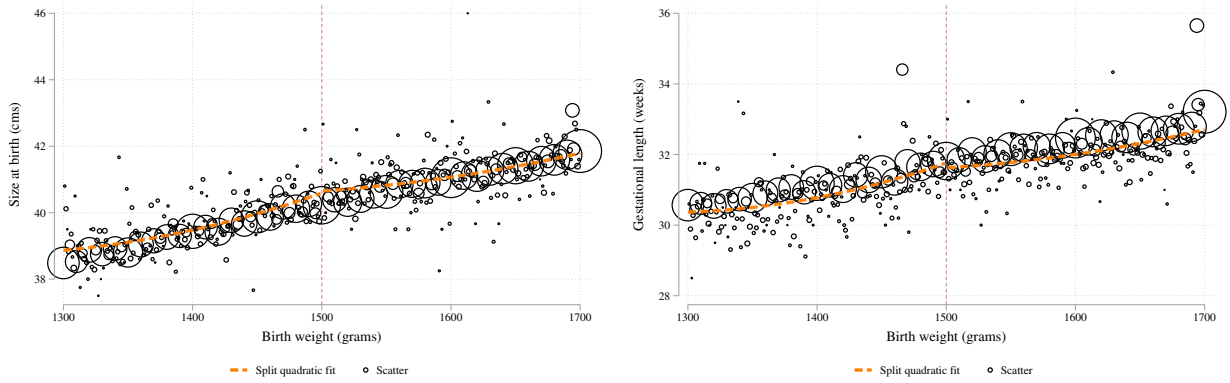
Figure A4: Birth Weight Frequency in Administrative Records



(a) Full Sample **(b)** 1300-1700 grams

Notes: Density plots are presented based on the full sample of births from 1992–2018. Panel (a) includes all births, while panel (b) limits only to births recorded as weighing between 1300 and 1700 grams (inclusive). In panel (b) 10 gram bins are plotted in the histogram.

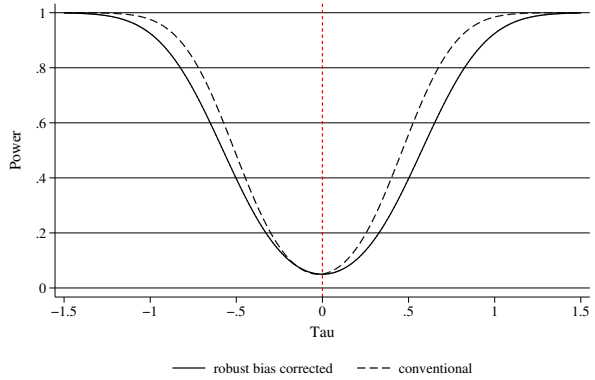
Figure A5: Observable Birth Outcomes by Birth Weight



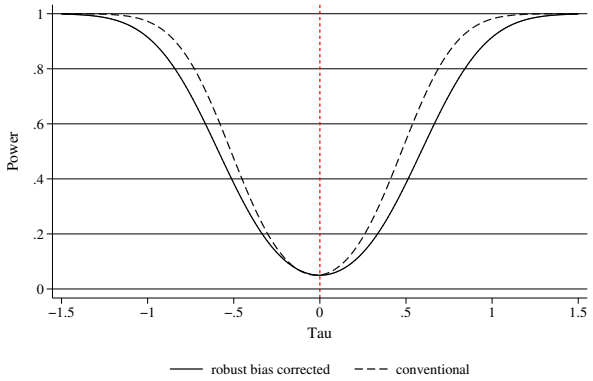
(a) Birth size (cms) **(b)** Gestational weeks

Notes: Scatter plots are presented where each point represents average birth size (panel (a)), or gestational weeks (panel (b)) in 50 gram birth weight bins. Each panel is based on the full sample of births from 1992–2018 between 1,300-1,700 grams, and point sizes reflect the relative frequency of the sample in each bin. Separate linear trends are plotted on each side of the 1,500 gram VLBW threshold.

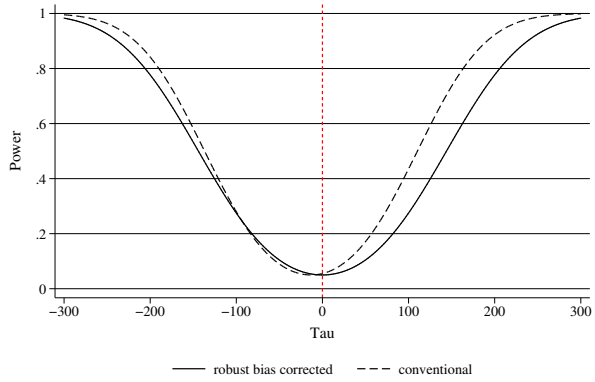
Figure A6: Power Analysis – RD Models of Intergenerational Impacts



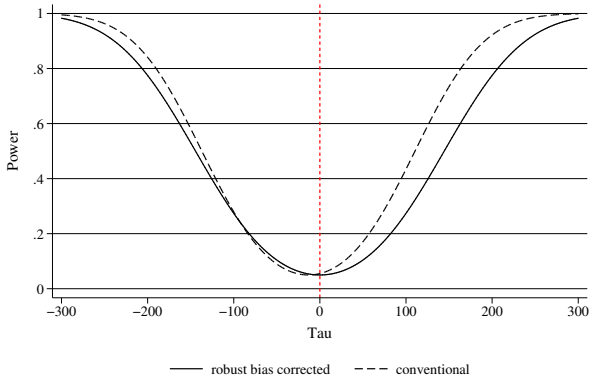
(a) Child's Gestational Length, Baseline



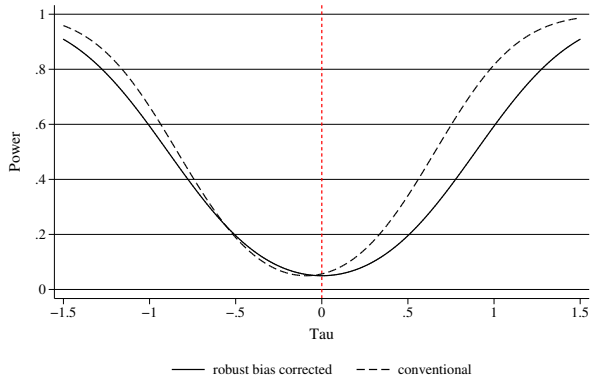
(b) Child's Gestational Length, Heaping Control



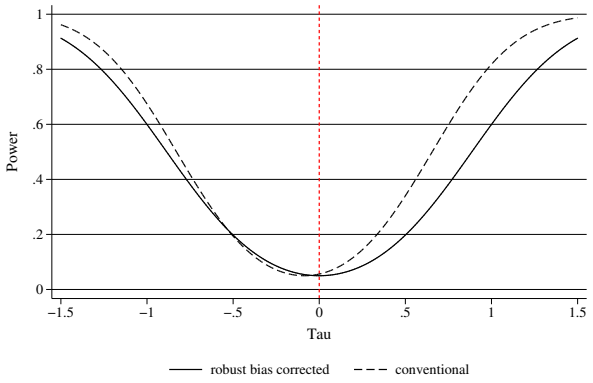
(c) Child's Birth weight, Baseline



(d) Child's Birth weight, Heaping Control



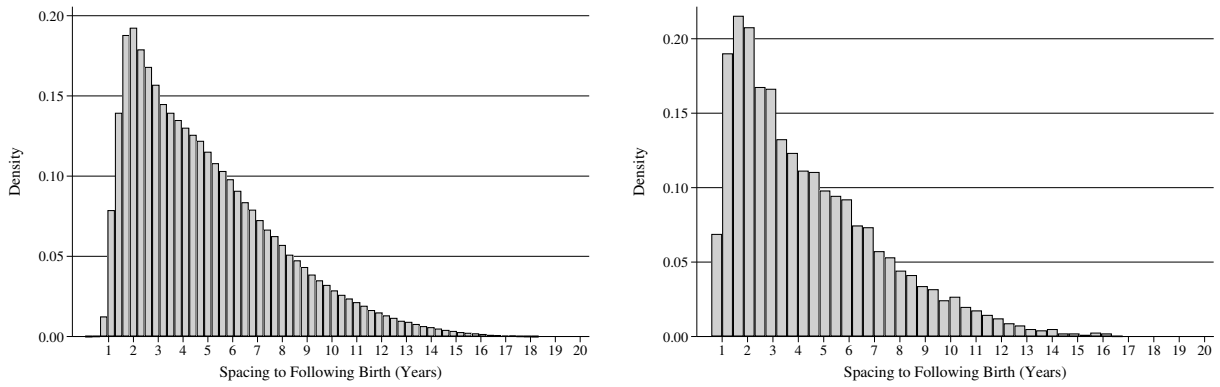
(e) Child's Birth length, Baseline



(f) Child's Birth length, Heaping Control

Notes: Curves in each figure plot the power of our regression discontinuity design to detect effect sizes indicated by Tau on the horizontal axes against a null of zero effects. Effects refer to intergenerational impacts, and as such, in each case, samples consist of second generation births corresponding to our principal estimation sample. Power cacluations follow Cattaneo et al. (2019b).

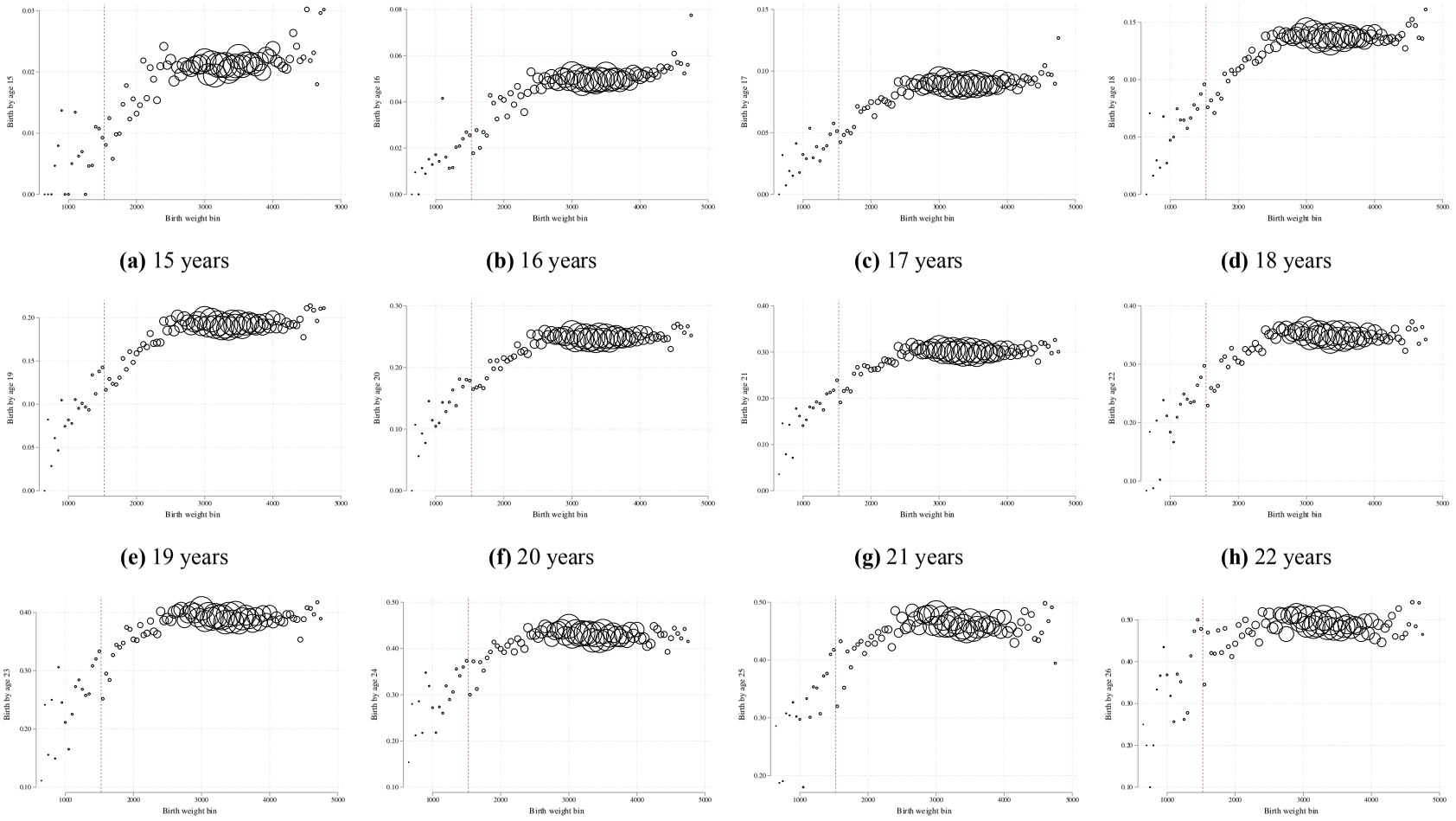
Figure A7: Observed Birth Spacing Between Children and their Future Siblings



(a) All Individuals **(b)** $1300 \leq$ Birth weight ≤ 1700

Notes: Histograms document the space in years between each observed birth and following births of each mother (provided following births are observed) in the full sample of administrative records from birth registries (panel (a)), and only those births with babies weighing between 1,300 and 1,700 grams (panel (b)).

Figure A8: Fertility and Health at Birth



(a) 15 years **(b)** 16 years **(c)** 17 years **(d)** 18 years

(e) 19 years **(f)** 20 years **(g)** 21 years **(h)** 22 years

(i) 23 years **(j)** 24 years **(k)** 25 years **(l)** 26 years

Notes: Refer to Notes to Figure 5. Identical plots are presented of the mean number of births by birth weight bins, however here for each age up to 26 years.

Table A4: Is there Policy-Driven Selection intro Childbirth? RD Estimates

	Mother			Father				Parent's	Observed	Urban
	Age	Education	Employed	Age	Education	Employed	Married	Age Diff.	Dad	Status
Birth weight < 1,500	-0.447	-0.708*	-0.002	-3.757***	-0.622	0.012	0.075*	-2.780**	-0.028	-0.087
	(0.318)	(0.409)	(0.069)	(0.906)	(0.386)	(0.051)	(0.040)	(1.085)	(0.070)	(0.060)
	[0.080]	[0.958]	[0.510]	[0.000]	[0.946]	[0.411]	[0.031]	[0.005]	[0.656]	[0.925]
Mean of Dep. Var.	19.910	11.228	0.183	23.373	11.492	0.853	0.072	3.326	0.842	0.874
Observations	431,165	430,709	371,405	363,913	363,640	313,001	430,572	363,913	431,167	431,167
Optimal Bandwidth	208.6	168.2	205.9	159.2	194.7	327.2	277.7	173.9	181.3	180.1
Effective Observations	1,016	786	886	635	752	1,295	1,411	690	859	859
Observations (left)	326	276	285	223	262	298	372	245	295	295
Observations (right)	690	510	601	412	490	997	1,039	445	564	564
q-sharpened p-value	0.223	0.201	0.417	0.001	0.208	0.377	0.198	0.050	0.348	0.223

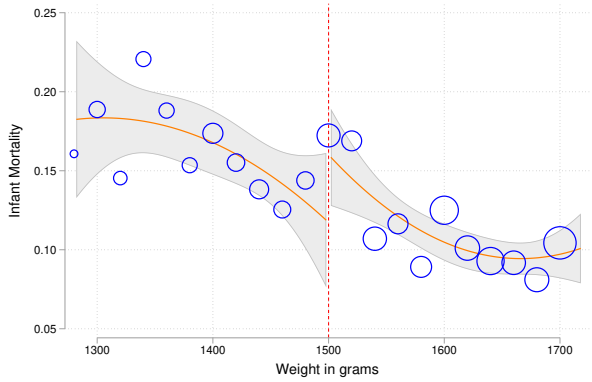
Notes: Each column displays estimates of the change in the given dependent variable from above to below the 1,500 gram assignment threshold for newborns. In each case, local linear regression is used with a triangular kernel, calculating the MSE optimal bandwidth of Calonico et al. (2020). Robust bias corrected standard errors clustered at the gram level are reported in parentheses. p-values for one-sided tests are shown in square brackets, where in each case the null hypothesis is that there is a negative policy impact. Significance stars for two-sided test: * p<0.10; ** p<0.05; *** p<0.01.

Table A5: Initial Policy Impacts on Infant Mortality Rates

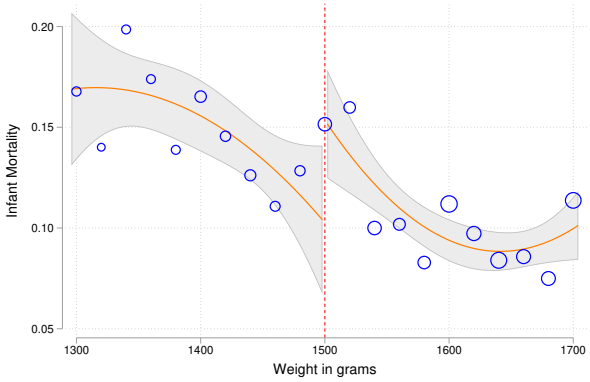
	Bharadwaj et al. (2013) Period		Full Period (1992-2018)		Early cohorts (1992-2006)	
	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	-0.0311**	-0.0165*	-0.0270***	-0.0207***	-0.0427***	-0.0253**
	[0.0123]	[0.0099]	[0.0080]	[0.0075]	[0.0142]	[0.0106]
Robust	-0.0131	-0.0189	-0.0295**	-0.0248***	-0.0111	-0.0280**
	[0.0147]	[0.0117]	[0.0120]	[0.0085]	[0.0169]	[0.0125]
Mean of Dep. Var.	0.120	0.111	0.102	0.098	0.121	0.112
Observations	3,913,016	3,913,016	6,525,050	6,525,050	3,675,886	3,675,886
Optimal Bandwidth	100.0	203.6	100.0	134.4	100.0	197.0
Effective Observations	5,340	12,843	9,464	13,815	5,000	10,095
Observations (left)	2,037	3,942	3,695	4,999	1,896	3,512
Observations (right)	3,303	8,901	5,769	8,816	3,104	7,473
Bharadwaj et al. (2013) Procedure	Y		Y		Y	
Optimal Bandwidth RBC		Y		Y		Y

Notes: Each column presents estimates of the impact of intensive treatment owing to the VLBW assignment on an indicator for infant mortality. Models in odd columns replicate the methods of Bharadwaj et al. (2013) using a 100 gram cut-off, and in even columns use the MSE optimal bandwidth of Calonico et al. (2020). In each case standard RD models are presented (top rows) which correspond to methods implemented by Bharadwaj et al. (2013), and robust bias corrected estimates and standard errors are presented below. Columns (1) and (2) use the exact same time period as that studied by Bharadwaj et al. (2013) (1992-2001), columns (3) and (4) consider the full sample of data used in this paper, and columns (5) and (6) consider only birth cohorts in which second generation mothers are born. Significance stars for two-sided test: * $p<0.10$; ** $p<0.05$; *** $p<0.01$.

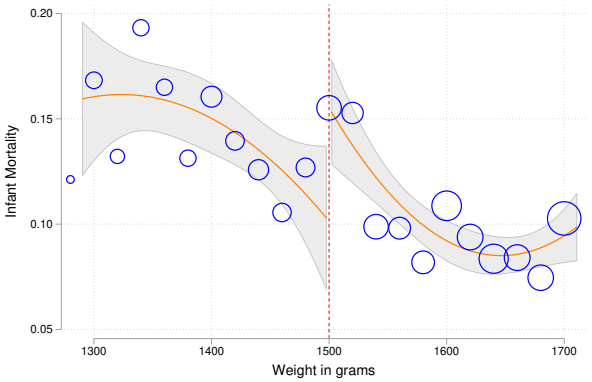
Figure A9: Birth weight Assignment Thresholds and Infant Mortality (Early Years Only)



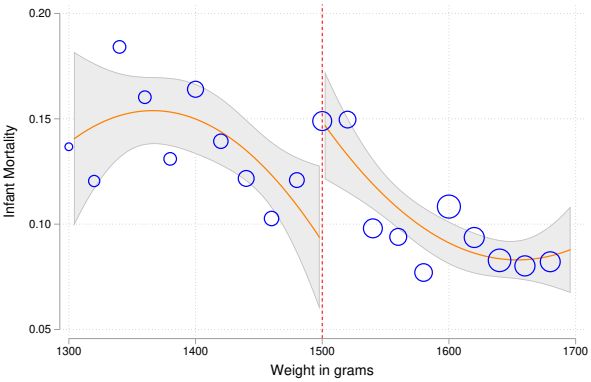
(a) $\leq$ 2001 and Optimal Bandwidth



(b) $\leq$ 2004 and Optimal Bandwidth



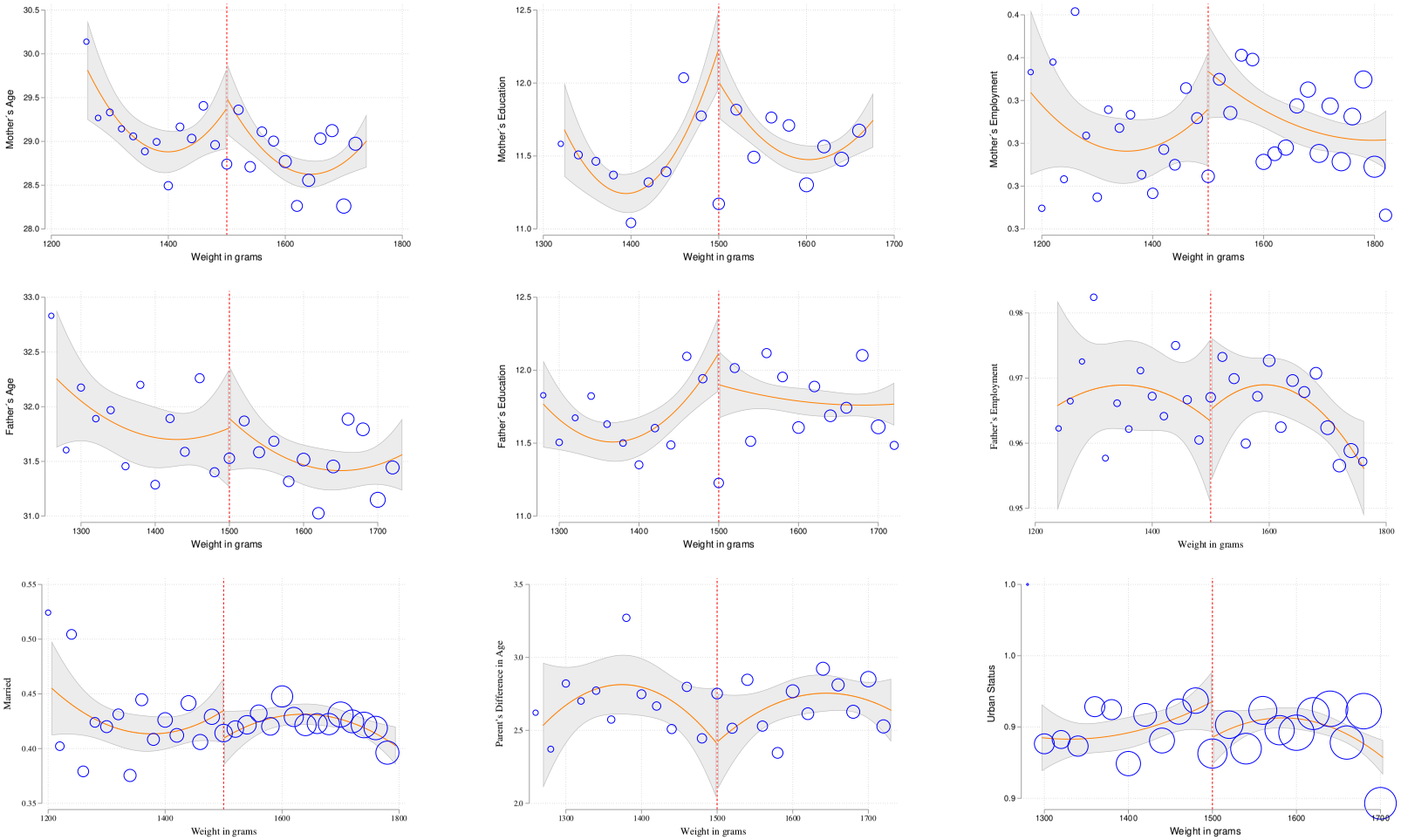
(c) $\leq$ 2005 and Optimal Bandwidth



(d) $\leq$ 2006 and Optimal Bandwidth

Notes: Refer to Notes to Figure 8. Identical plots are presented to panel (c), however here documenting alternative considerations of "early cohorts", as indicated in captions.

Figure A10: Balance Tests – Generation 1 Family and Birth Characteristics



Notes: Plots document balance tests examining characteristics of individuals within the Calonico et al. (2020) RBC optimal bandwith of the 1,500 gram treatment threshold. Binned averages for each outcome are presented in 20 gram birth weight bins, with circle sizes documenting the relative proportion of individuals in each bin. A separate split quadratic and confidence intervals are documented on each side of the cut-off. Formal tests of discontinuities are presented in Table A6.

Table A6: Balance Tests – Generation 1 Family and Birth Characteristics

	Mother			Father				Parent's	Observed	Urban
	Age	Education	Employed	Age	Education	Employed	Married	Age Diff.	Dad	Status
Birth weight < 1,500	-0.1345	0.1696	-0.0313*	-0.1220	0.1081	-0.0069	0.0048	-0.0800	-0.0126	0.0037
	[0.2716]	[0.2227]	[0.0190]	[0.4152]	[0.2421]	[0.0046]	[0.0169]	[0.2336]	[0.0085]	[0.0106]
Mean of Dep. Var.	28.875	11.527	0.326	31.590	11.726	0.964	0.425	2.621	0.9555	0.897
Observations	6,104,565	6,098,233	6,105,345	5,835,120	5,466,070	4,893,004	6,105,345	5,827,941	6,097,490	6,105,345
Optimal Bandwidth	239.7	176.9	321.8	233.7	220.8	262.1	219.5	230.5	173.6	203.3
Effective Observations	24,502	17,189	36,367	22,976	20,210	22,063	22,136	22,790	17,012	20,983
Observations (left)	10,174	10,101	10,175	9,645	8,884	7,831	10,175	9,572	10,100	10,175
Observations (right)	6,094,391	6,088,132	6,095,170	5,825,475	5,457,186	4,885,173	6,095,170	5,818,369	6,087,390	6,095,170

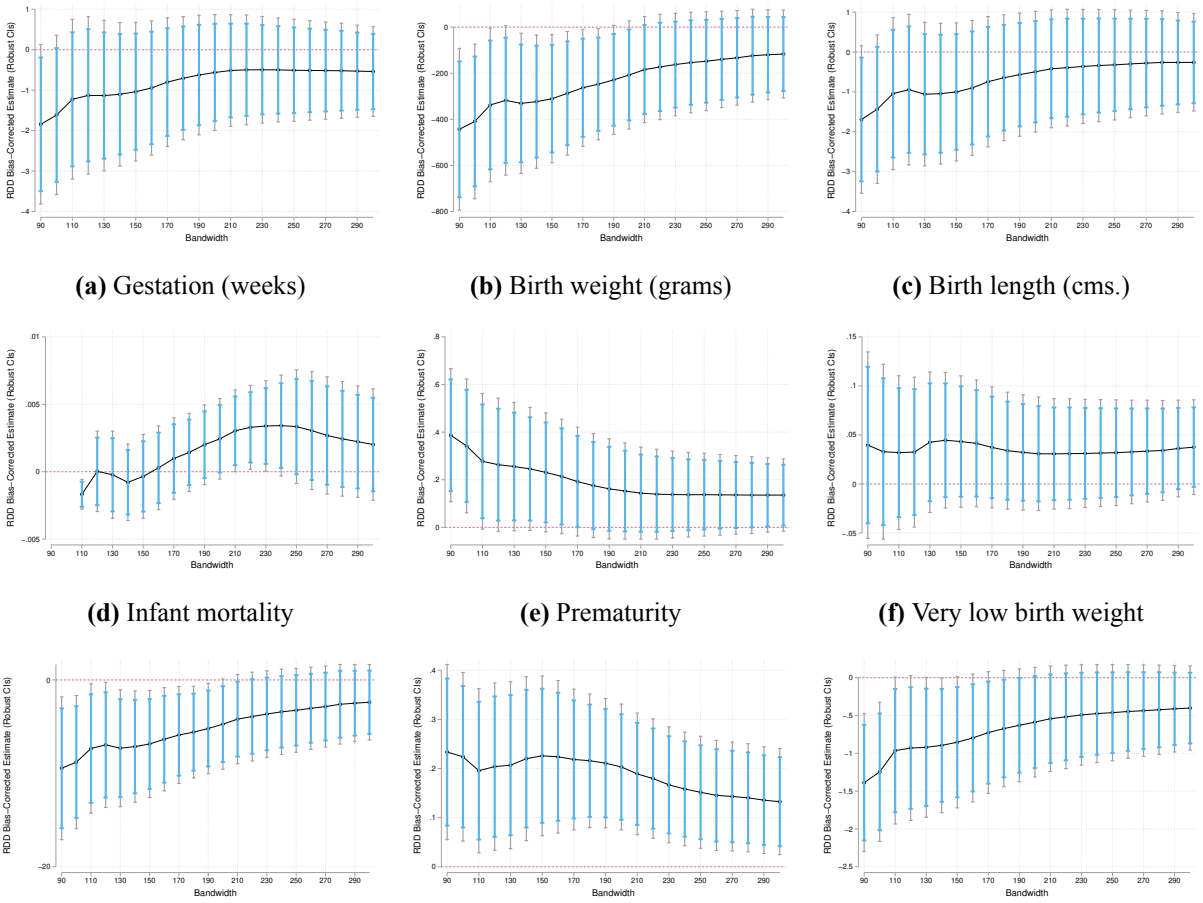
Notes: Each column displays estimates of the change in the given dependent variable from above to below the 1,500 gram assignment threshold for newborns. Each outcome refers to mother, father, or family level characteristics of all first generation observations, and as such are balance tests examining assignment to treatment. In each case, local linear regression is used with a triangular kernel, calculating the MSE optimal bandwidth of Calonico et al. (2020). Robust bias corrected standard errors clustered at the gram level are reported in parentheses. Significance stars for two-sided test: * $p<0.10$; ** $p<0.05$; *** $p<0.01$.

Table A7: Testing for birth heaping at the cut-off

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Total Difference										
< 1,500 gram threshold	9.500	2.950	-0.367	-0.225	-2.320	-5.627	-11.330	-14.368	-18.860	-27.230
	[81.433]	[64.140]	[51.129]	[44.876]	[44.857]	[34.144]	[33.759]	[29.062]	[27.359]	[24.457]
Total Births	1,609	3,497	5,249	7,217	9,632	13,394	19,695	23,626	29,401	39,918
Gram-Specific cells	20	40	60	80	100	150	200	250	300	400
R-Squared	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003
Panel B: Total Difference and Distance Trend										
< 1,500 gram threshold	-3.867	12.921	13.533	9.115	9.249	6.769	10.936	7.948	9.045	10.963
	[160.156]	[131.198]	[104.887]	[91.570]	[90.928]	[69.377]	[68.090]	[58.647]	[55.069]	[49.111]
Absolute Difference (grams)	29.618	10.253	4.582	2.841	2.579	0.566	0.936	0.417	0.472	0.401
	[18.251]	[7.744]	[4.178]	[2.752]	[2.194]	[1.122]	[0.828]	[0.571]	[0.447]	[0.300]
Absolute Difference (grams) ×	2.430	-0.950	-0.897	-0.456	-0.454	-0.326	-0.441	-0.354	-0.370	-0.380
<1,500 grams	[25.811]	[10.952]	[5.908]	[3.892]	[3.103]	[1.586]	[1.171]	[0.808]	[0.633]	[0.424]
Total Births	1,609	3,497	5,249	7,217	9,632	13,394	19,695	23,626	29,401	39,918
Gram-Specific cells	20	40	60	80	100	150	200	250	300	400
R-Squared	0.264	0.082	0.034	0.023	0.024	0.002	0.009	0.003	0.006	0.008
Absolute margin (grams)	10	20	30	40	50	75	100	125	150	200

Notes: Each column presents a regression of the number of births observed on an indicator for whether the recorded weight is below 1,500 grams, in a fixed range around the 1,500 gram cut-off. This range is varied across columns from± 20 grams to ± 400 grams, as indicated in the table footer. In panel (A) a single binary "< 1,500 gram" indicator is included, while in panel (B) this is additionally interacted with a control for absolute difference in grams, to control for the fact that the number of births gradually increase as birth weight increases. Significance stars for two-sided test: * p<0.10; ** p<0.05; *** p<0.01.

Figure A11: RD Estimates of Intergenerational Impacts Varying Estimation Bandwidths



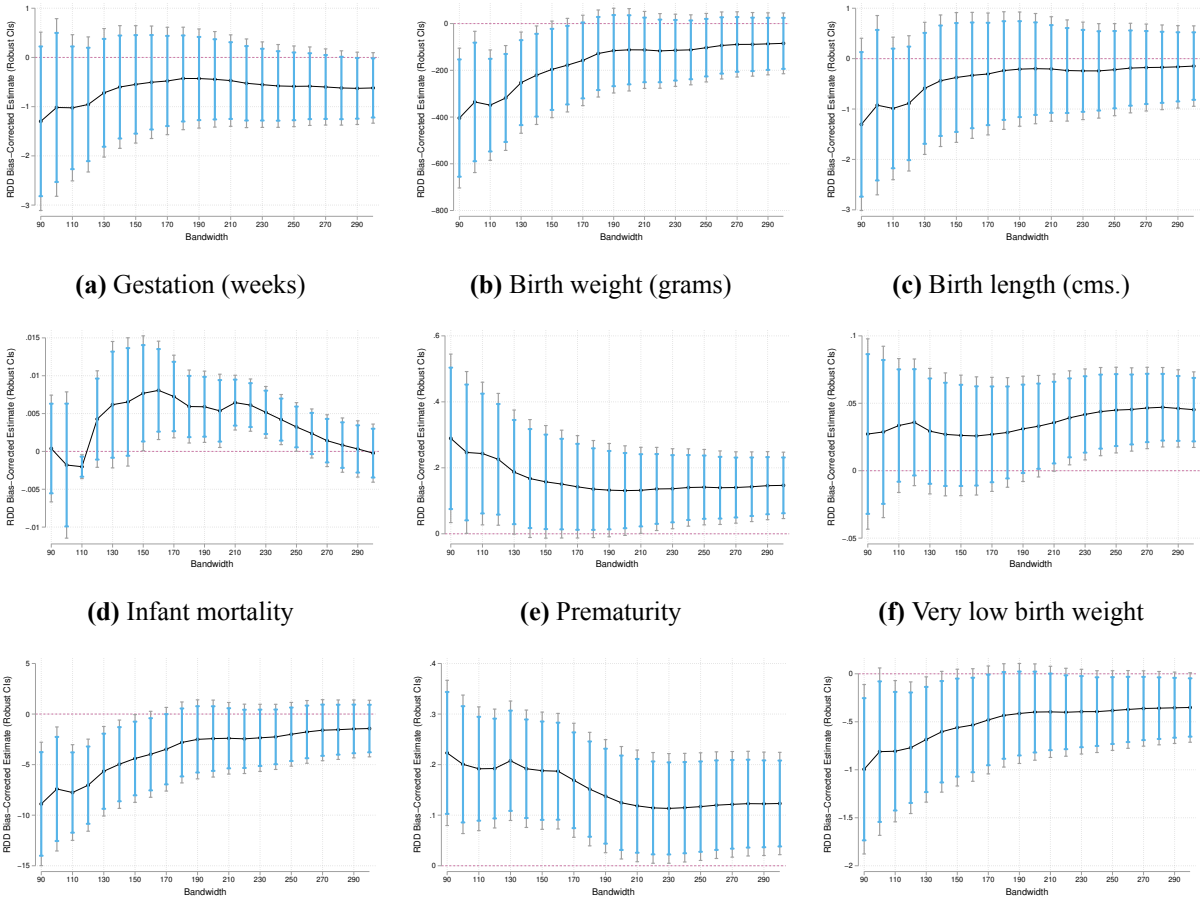
(a) Gestation (weeks) **(b)** Birth weight (grams) **(c)** Birth length (cms.)

(d) Infant mortality **(e)** Prematurity **(f)** Very low birth weight

(g) Fetal growth rate **(h)** Sex ratio **(i)** Anderson index

Notes: Each sub-plot portrays the impact of crossing the VLBW threshold on a specific second-generation health measure from Table 2, where the RDD specification is estimated within a bandwidth manually set to 90,100,110,…,300 grams. Point estimates and 90% and 95% confidence intervals (blue and grey lines respectively) are shown using bias corrected estimates with heaping controls. The estimation sample consists of all children of first generation mothers who were born at 32 weeks or greater of gestation.

Figure A12: RD Estimates of Intergenerational Impacts Varying Estimation Bandwidths with Constant Relative Bias



(a) Gestation (weeks) **(b)** Birth weight (grams) **(c)** Birth length (cms.)

(d) Infant mortality **(e)** Prematurity **(f)** Very low birth weight

(g) Fetal growth rate **(h)** Sex ratio **(i)** Anderson index

Notes: Each sub-plot portrays the impact of crossing the VLBW threshold on a specific second-generation health measure from Table 2, where the RDD specification is estimated within a bandwidth manually set to 90,100,110,…,300 grams. Point estimates and 90% and 95% confidence intervals (blue and grey lines respectively) are shown using bias corrected estimates with heaping controls, where the bias correction calculation is conducted by holding a constant relative bias correction range to optimal bandwidth range as that used in principal models displayed in the paper. The estimation sample consists of all children of first generation mothers who were born at 32 weeks or greater of gestation.

Table A8: Alternative Specifications – Gestational weeks, birth weight, birth length, infant mortality

	Linear Running Variable				Quadratic Running Variable			
	Main	CERRD	Baseline	Bias Corrected	Main	CERRD	Baseline	Bias Corrected
Panel A: Gestations weeks								
Birth weight < 1500	–0.441	–0.565	–0.353	–0.441	–0.689	–0.470	–0.652	–0.689
	(0.502)	(0.533)	(0.407)	(0.407)	(0.593)	(0.659)	(0.491)	(0.491)
Observations	421,403	421,403	421,403	421,403	421,403	421,403	421,403	421,403
Optimal Bandwith	196.6	136.4	196.6	196.6	355.1	233.8	355.1	355.1
Panel B: Birth weight (grams)								
Birth weight < 1500	–148.907	–264.845***	–118.088	–148.907*	–203.071*	–333.865***	–170.881	–203.071*
	(96.771)	(95.014)	(80.459)	(80.459)	(118.110)	(125.898)	(105.365)	(105.365)
Observations	421,382	421,382	421,382	421,382	421,382	421,382	421,382	421,382
Optimal Bandwith	171.2	118.8	171.2	171.2	256.9	169.1	256.9	256.9
Panel C: Birth Length (cms)								
Birth weight < 1500	–0.199	–0.357	–0.087	–0.199	–0.426	–0.702	–0.309	–0.426
	(0.563)	(0.559)	(0.464)	(0.464)	(0.726)	(0.763)	(0.620)	(0.620)
Observations	421,376	421,376	421,376	421,376	421,376	421,376	421,376	421,376
Optimal Bandwith	193.5	134.2	193.5	193.5	300.0	197.5	300.0	300.0
Panel D: Infant Mortality								
Birth weight < 1500	0.007*	–0.000	0.002*	0.007***	0.002	–0.009***	–0.001	0.002
	(0.004)	(.)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Observations	421,796	421,796	421,796	421,796	421,796	421,796	421,796	421,796
Optimal Bandwith	142.3	98.7	142.3	142.3	188.4	124.0	188.4	188.4

Notes: Alternative specifications are documented for each inter-generational outcome considered in Table 2. Each panel considers a specific outcome, based on local linear (columns 1-4), and local quadratic models (columns 5-8). Alternative models consist of robust bias corrected models (column 1 and 5), models using Calonico et al. (2020)'s minimum error coverage bandwidth selection (columns 2 and 4), using standard RD models without robust bias correction (columns 3 and 5), and using bias corrected but not robust models (columns 4 and 8). Significance stars for two-sided test: * $p<0.10$; ** $p<0.05$; *** $p<0.01$.

Table A9: Alternative Specifications – Prematurity, VLBW, Fetal Growth, Sex ratio, Anderson Index

	Linear Running Variable				Quadratic Running Variable			
	Main	CERRD	Baseline	Bias Corrected	Main	CERRD	Baseline	Bias Corrected
Panel E: Prematurity								
Birth weight < 1500	0.132*	0.169**	0.118**	0.132**	0.144*	0.156*	0.142**	0.144**
	(0.072)	(0.078)	(0.059)	(0.059)	(0.081)	(0.092)	(0.068)	(0.068)
Observations	421,796	421,796	421,796	421,796	421,796	421,796	421,796	421,796
Optimal Bandwith	189.8	131.7	189.8	189.8	350.4	230.7	350.4	350.4
Panel F: Very low birth weight status								
Birth weight < 1500	0.045***	0.030*	0.042***	0.045***	0.022	0.033	0.026	0.022
	(0.016)	(0.018)	(0.013)	(0.013)	(0.030)	(0.030)	(0.026)	(0.026)
Observations	421,796	421,796	421,796	421,796	421,796	421,796	421,796	421,796
Optimal Bandwith	257.7	178.7	257.7	257.7	270.4	178.0	270.4	270.4
Panel G: Fetal growth rate								
Birth weight < 1500	–2.736	–5.072***	–2.232	–2.736*	–4.580*	–7.823***	–3.959*	–4.580**
	(2.010)	(1.898)	(1.663)	(1.663)	(2.539)	(2.661)	(2.248)	(2.248)
Observations	421,382	421,382	421,382	421,382	421,382	421,382	421,382	421,382
Optimal Bandwith	178.2	123.6	178.2	178.2	247.0	162.6	247.0	247.0
Panel H: Sex ratio								
Birth weight < 1500	0.121**	0.179***	0.100**	0.121**	0.183***	0.237***	0.166***	0.183***
	(0.057)	(0.053)	(0.049)	(0.049)	(0.062)	(0.071)	(0.058)	(0.058)
Observations	421,762	421,762	421,762	421,762	421,762	421,762	421,762	421,762
Optimal Bandwith	234.4	162.6	234.4	234.4	257.1	169.2	257.1	257.1
Panel I: Anderson Index								
Birth weight < 1500	–0.393*	–0.464*	–0.339*	–0.393**	–0.494	–0.668*	–0.449*	–0.494*
	(0.220)	(0.246)	(0.180)	(0.180)	(0.313)	(0.343)	(0.264)	(0.264)
Observations	421,796	421,796	421,796	421,796	421,796	421,796	421,796	421,796
Optimal Bandwith	232.2	161.1	232.2	232.2	327.3	215.5	327.3	327.3

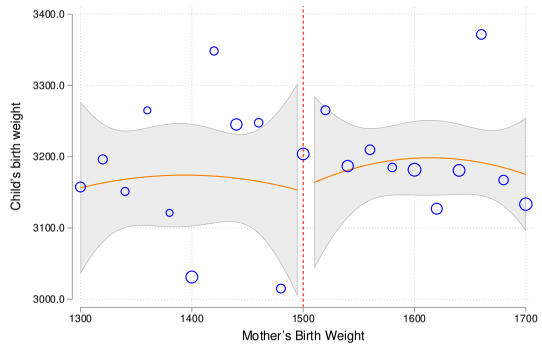
Notes: Refer to notes to Table A8. Identical models are presented, however considering alternative outcomes in each of panels E-I. Significance stars for two-sided test: * p<0.10; ** p<0.05; *** p<0.01.

Table A10: Intensive Health Investments and Birth Outcomes of the Second Generation (All births)

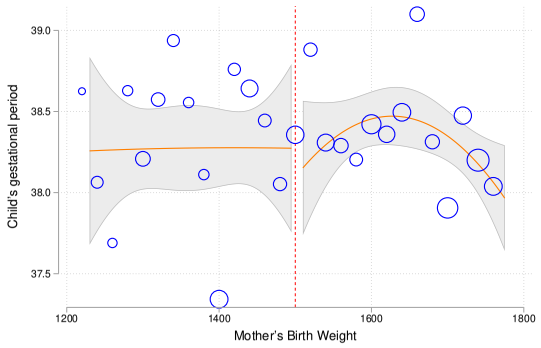
Panel A: Baseline Variables	Gestation (weeks) (1)	Birth weight (grams) (2)	Birth length (cms) (3)	Infant Mortality (4)	
Birth weight < 1,500	-0.054	-79.295	-0.065	-0.011**	
	(0.243)	(63.712)	(0.320)	(0.005)	
	[0.589]	[0.893]	[0.581]	[0.017]	
Mean of Dep. Var.	38.318	3184.247	48.649	0.008	
Observations	423,365	423,344	423,337	423,760	
Optimal Bandwidth	273.2	201.5	250.0	254.5	
Effective Observations	2,037	1,529	1,903	1,907	
Observations (left)	733	608	703	704	
Observations (right)	1,304	921	1,200	1,203	
q-sharpened p-value	0.398	0.166	0.398	0.074	
Panel B: Transformed Measures	Prematurity (5)	Very low birth weight (6)	Fetal growth rate (7)	Sex ratio (8)	Anderson Index (9)
Birth weight < 1,500	0.080*	0.024***	-2.061	0.108***	-0.184
	(0.043)	(0.009)	(1.492)	(0.038)	(0.141)
	[0.969]	[0.996]	[0.916]	[0.998]	[0.904]
Mean of Dep. Var.	0.091	0.013	82.617	0.500	-0.054
Observations	423,760	423,760	423,344	423,726	423,760
Optimal Bandwidth	174.0	215.7	204.6	194.5	234.0
Effective Observations	1,257	1,565	1,529	1,364	1,700
Observations (left)	509	619	608	555	657
Observations (right)	748	946	921	809	1,043
q-sharpened p-value	0.101	0.030	0.155	0.030	

Notes: Each column displays estimates of the change in the given dependent variable from above to below the 1,500 gram assignment threshold for mothers. In each case, local linear regression is used with a triangular kernel, calculating the MSE optimal bandwidth of Calonico et al. (2020). Robust bias corrected standard errors are reported in parentheses. Below standard errors, a one tailed t-test is calculated, which can be viewed as the support in favour of there actually being *positive* intergenerational transmission to the second generation. q-sharpened p-values refer to corrections conducted across the entire class of outcomes. * p<0.10; ** p<0.05; *** p<0.01.

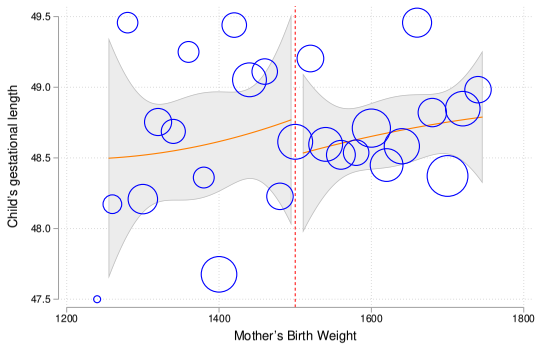
Figure A13: Descriptive Plots of Parental Policy Receipt and Child Health Measures (All births)



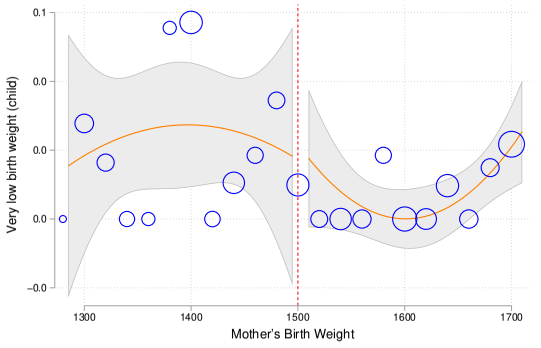
(a) Birth weight



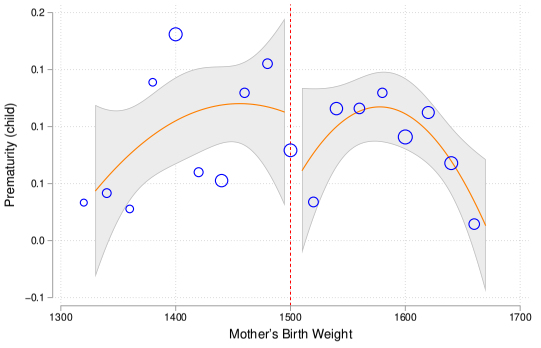
(b) Gestational period



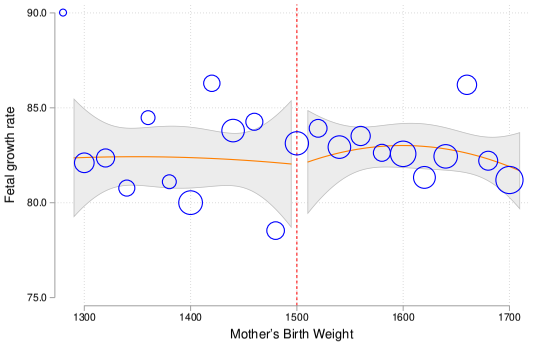
(c) Gestational length



(d) Very low birth weight



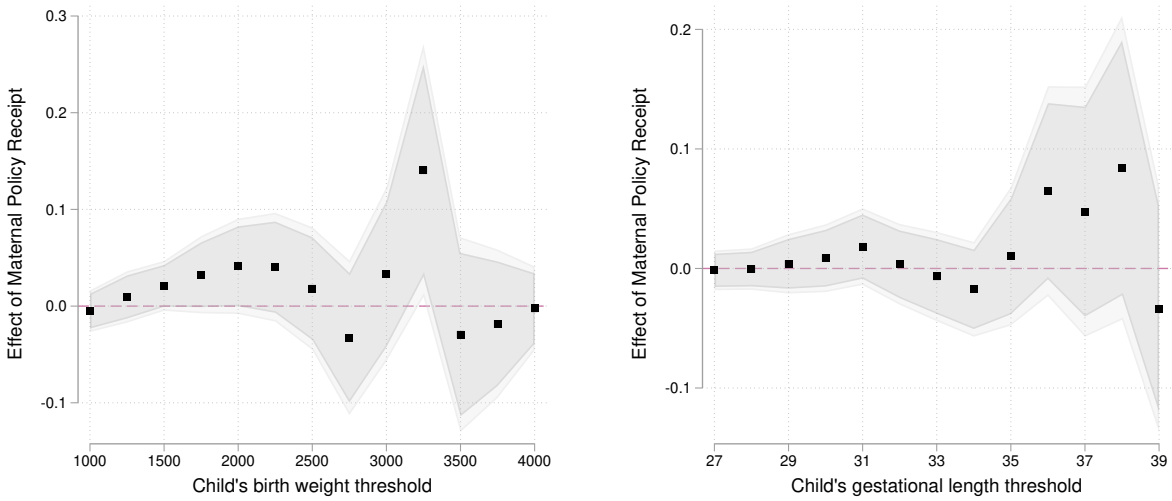
(e) Prematurity



(f) Fetal growth rate

Notes: Plots show separate quadratic fits estimated on each side of the 1500 gram cut-off, in each case restricting attention to observations within the optimal bandwidth following Calonico et al. (2014). 95% confidence intervals of the quadratic fit are estimated, and circles represent average outcomes in 20 gram bins. Observations at 1500 grams are plotted in binned averages, but are not used in estimating the quadratic fit.

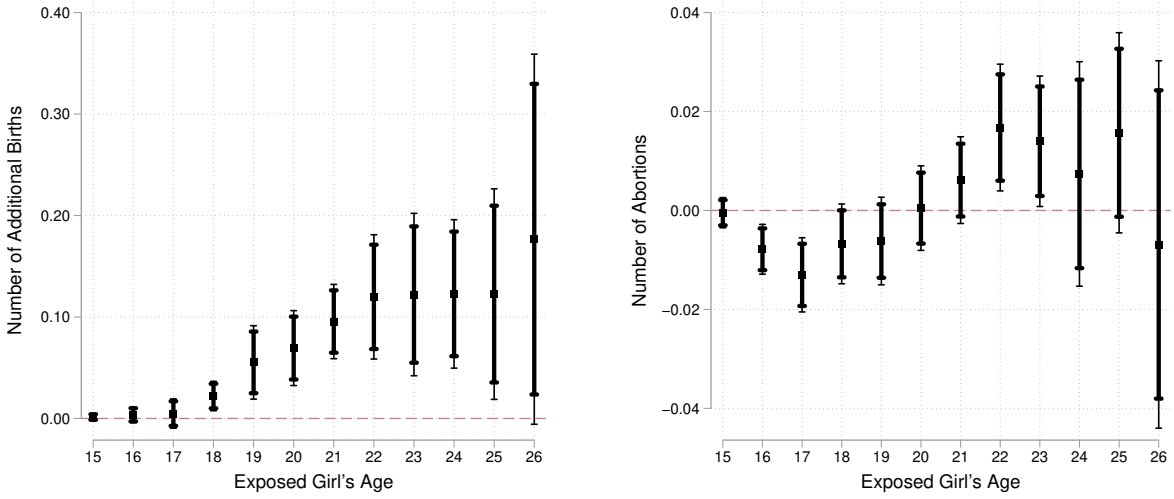
Figure A14: Distributional Impacts of Early Life Health Interventions on Second Generation Health Stocks (All births)



(a) Birth weight (All) **(b)** Gestational length (All)

Notes: Each point estimate (black square) and 90 and 95% confidence interval (dark and light shaded areas respectively) refers to RDD estimates of the likelihood that a birth occurring to a treated girl born has health stocks (birth or gestational weight) below the cut-off indicated on the x-axis of each sub-plot. Panels (a) and (c) refer to all individuals, while panel (b) and (d) refer to first generation individuals who were born at above 32 weeks of gestation, and are hence more clearly targeted by the policy.

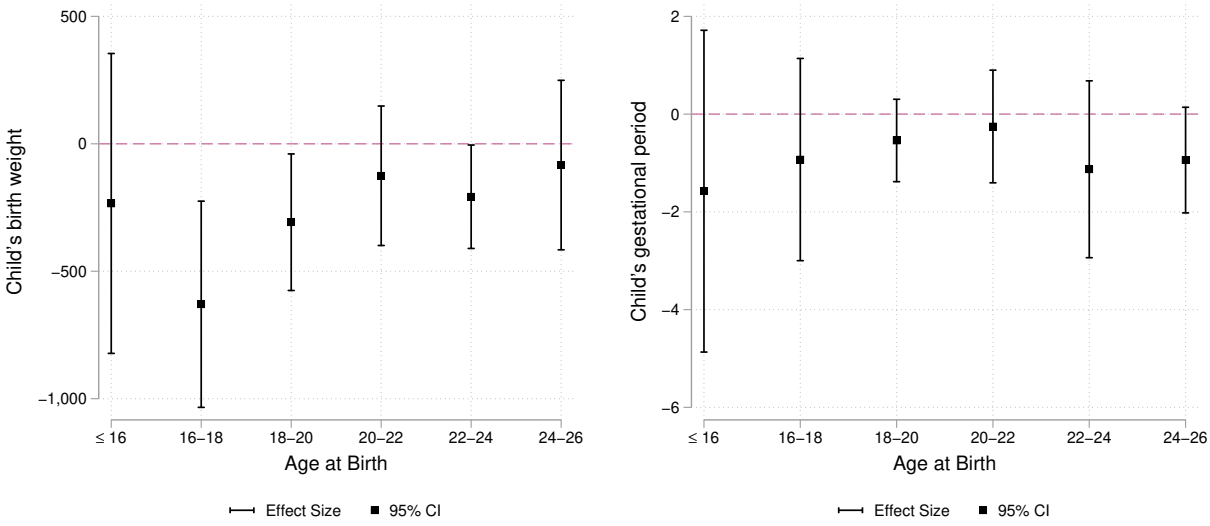
Figure A15: Impacts of Early Life Health Interventions on Fertility and Spontaneous Abortions (All births)



(a) Fertility **(b)** Abortions

Notes: Each point estimate and confidence interval refer to the impacts of early life medical investment on the number of births an individual has had by each age (panel (a)), and the number of abortions observed in hospitalization data (panel (b)). Thicker black error bars present 90% CIs, while thinner error bars report 95% CIs. All estimates follow the procedures laid out in section 4, and report RBC estimates using a local linear regression with a triangular kernel in the MSE optimal bandwidth.

Figure A16: Estimated RDD Effects by Second Generation Mother's Age



(a) Birth weight

(b) Gestational length

Notes: Each point estimate (black square) and 95% confidence interval refer to RDD estimates of the effect of treatment on birth weight (panel (a)) and gestational length (panel (b)), for second generation mothers are ages indicated in panels. Each estimate is generated using the same sample and methods described in Notes to Table 2, however sub-setting exclusively to the ages indicated on the horizontal axis.

B An Extended Model with Dynamic Health Stocks

Here we note how the model laid out in section 4 of the paper generalizes to a more flexible Grossman (2000)-style health model, where individual health is measured as a dynamic flow, depending on health at birth, as well as health in previous periods. Consider health at adult-hood, when women are at risk of falling pregnant, and when one wishes to consider the interpretation of the RD model laid out in the paper, examining the inter-generational transmission of health at birth. In this case, the analogue of health stocks of generation 1 laid out in the paper is:

$$H_{it} = \phi_t H_i + \mu_t H_{i,t-1}(H_i, D_i, I_t^{post}, \zeta_i) + \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, H_{i,t-1}, \zeta_i) + X_{it}\beta_t + \upsilon_{it}, \quad (6)$$

where now, H_{it} is explicitly a function of $H_{i,t-1}$, and in turn, depends recursively on health in all previous periods: $H_{it} \equiv H_{it}(H_{i,t-1}, H_i, D_i, I_t^{post}, X)$. We note two specific relevant considerations of this model of health stocks. Firstly, it makes explicit that like parental investments, previous health stocks may depend upon initial health, treatment at birth, and subsequent stocks, as well as parental investment. Importantly, this implies that H_{it-1} may shift owing to threshold crossing, which must be taken into account when considering the interpretation of a reduced form RDD parameter during adulthood. Secondly, we note that health stocks at birth are explicitly indicated as H_i, to note the lack of dependency on these response variables.

Now, as laid out in the paper, equations 2-5 capture the implications of selection into fertility on the estimated intergenerational impacts of health investment at birth. To see how this impacts the final decomposition of the estimated RDD parameter $\widehat{\alpha}$, we generalize the model here to incorporate $H_{i,t-1}$. First, note that latent fertility now depends upon accumulated health stocks:

$$fert_{it}^* = \gamma_1 H_i + \gamma_2 D_i + \gamma_{3t} H_{i,t-1}(H_i, D_i, I_t^{post}, \zeta_i) + \gamma_{4t} I_t^{post}(H_i, D_i, H_{i,t-1}, \zeta_i) + X_{it}\pi_t + \iota_{it}. \quad (7)$$

Similarly, the model of birth weight of generation 2 also depends upon the mother's accumulated health stocks:

$$BW_{ij} = H_{ij} + \phi_t H_i + \mu_t H_{i,t-1}(H_i, D_i, I_t^{post}, \zeta_i) + \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, H_{i,t-1}, \zeta_i) + X_{it}\beta_t + u_{ij}, \quad (8)$$

with BW_{ij} observed for individuals only if $fert_{it} = 1$. Following the model in the paper, the conditional expectation of birth weight is now:

$$\begin{aligned} E(BW_{ij}|fert_{it} = 1) \quad = \quad & H_{ij} + \mu_t H_{i,t-1}(H_i, D_i, I_t^{post}, \zeta_i) + \phi_t H_i + \\ & \psi_t D_i + \varphi_t I_t^{post}(H_i, D_i, H_{i,t-1}, \zeta_i) + X_{it}\beta_t + \\ & \rho\sigma_u \lambda[H_i, H_{i,t-1}(H_i, D_i, I_t^{post}, \zeta_i), D_i, I^{post}(H_i, D_i, \zeta_i)] \end{aligned} \quad (9)$$

where all notation follows section 4 of the paper. Finally, in this case, the extended interpretation of the RDD estimate $\widehat{\alpha}$ is now the following:

$$\widehat{\alpha} = \psi_t \cdot \kappa + \mu_t \cdot \Delta H_{it-1}(c) + \varphi_t \cdot \Delta I_t^{post}(c) + \rho\sigma_u \Delta\lambda[c, \Delta H_{it-1}(c), \Delta I_t^{post}(c)]. \quad (10)$$

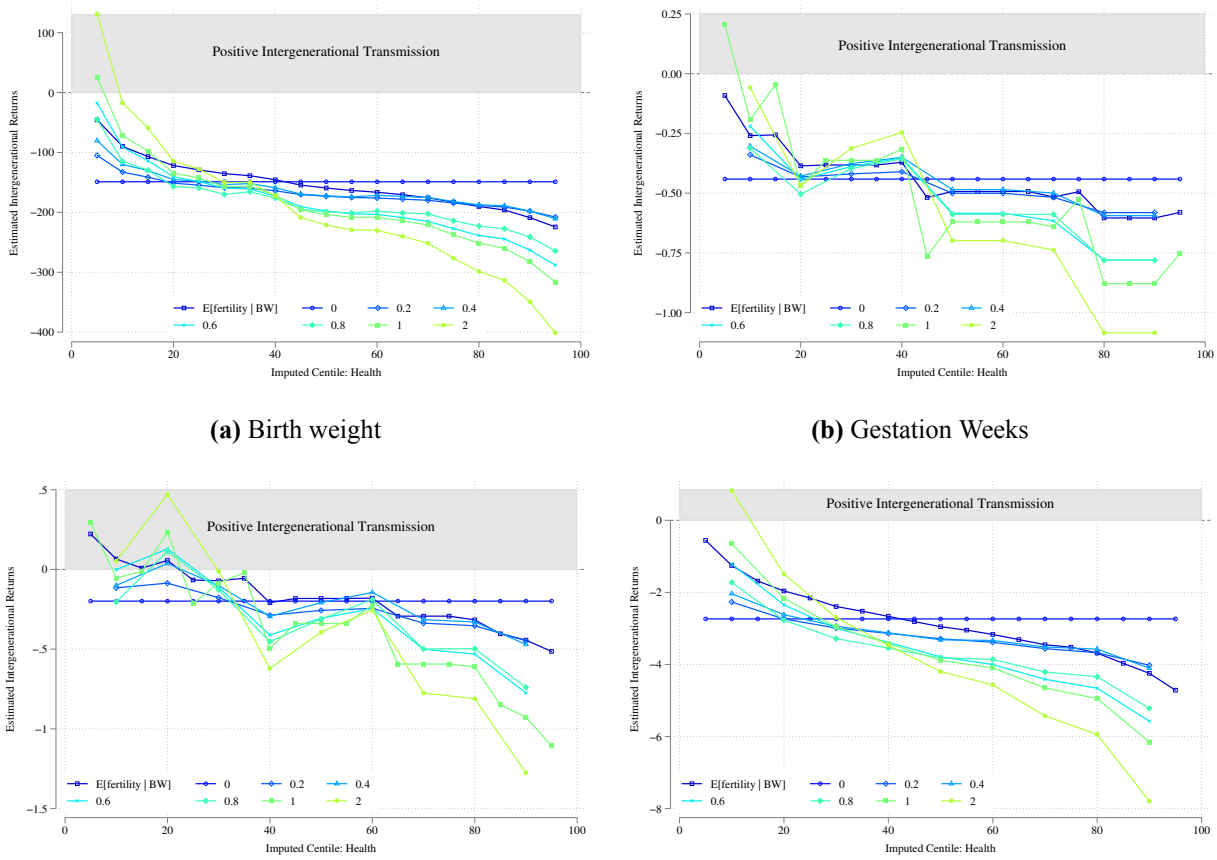
There are two upshots from this extended model allowing for dynamics in mother's health. Firstly, this opens up a direct new channel which may partially explain the RD estimate, which is that mother's health at $t-1$ (and recursively before that) impacts child's health at t, given that health investments at birth may directly impact mother's health at time t. This is indicated by $\mu_t \cdot \Delta H_{it-1}(c)$. Secondly, it additionally opens up a new sub-channel within the inverse Mills ratio, which is that a mother's health stock at time $t-1$ makes her more or less likely to give birth. This is indicated by $\Delta H_{i,t-1}(c)$ within the inverse Mills ratio, where this is relevant if treatment receipt shifts later-life health around the cutoff c. Note then that we can 'close off' this channel if we can show that women are not more or less healthy at each time t given treatment receipt.

C Two Dimensional Selection Counterfactual (Fertility and Health)

In Section 5.2.3 of the paper we examine whether survival selection in the first generation can explain observed results in the second generation. This analysis considers counterfactual stocks of children who did not exist in the second generation due to selective survival in the first generation. More generally, it is illustrative to consider the range of counterfactual scenarios over which it would be sufficient to turn negative observed intergenerational impacts into positive effects. In Table 4 we assume that all imputed individuals would have had fertility rates which were equal to those of surviving individuals in their birth weight bin. However, these individuals may have had quite different fertility profiles had they survived. Thus, in Figure C1 we consider a broader range of counterfactual scenarios. We plot RD estimates where counterfactual assumptions for health are described on the horizontal axis, and assumed fertility (the average number of births up to the end of the observed data) is varied across line plots. These counterfactual fertility assumptions range from 0 births per imputed individual (in which case the RD sample will remain unchanged) up to 2 additional births per imputed individual. Counterfactual health outcomes are allowed to range from the 5^{th} percentile up to the 95^{th} percentile of observed outcomes. Areas shaded in grey are consistent with positive intergenerational transfers.

The results from this activity are plotted for birth weight (Figure C1a), gestational weeks (Figure C1b), size at birth (Figure C1c), and fetal growth rate (Figure C1d). Binary outcomes are not considered, as these often entirely omit extreme outcomes such as VLBW which occur in less than 5% of cases. Considering birth weight as the outcome of interest, it appears highly unlikely that selective survival at birth could explain the observed negative intergenerational transfers from mothers to children. For this to be the case, non-surviving individuals would have had to be highly fertile (each having 1 birth), while at the same time giving birth to children with very low health stocks (all at the 5^{th} health percentile). Similar extreme patterns are observed for both gestational weeks and the fetal growth rate. In the case of size at birth, more feasible counterfactual outcomes exist which could explain away the negative intergenerational transmission (for example, average birth rates, and all counterfactual babies being born at the 20^{th} health percentile or below. Nevertheless, across all outcomes considered, the evidence broadly suggests that correcting for survival at birth in the first generation would not be sufficient to turn around observed results in all health dimensions.

Figure C1: Inter-generational Transmission Under Alternative Health and Fertility Counterfactuals



(a) Birth weight

(b) Gestation Weeks

(c) Size at birth

(d) Fetal growth rate

Notes: Each point represents estimates of the impact of intensive health receipt at first generation births on the second generation health outcomes indicated as plot labels. Each estimate comes from a separate RD model based on all surviving observations, as well as imputed outcomes for individuals estimated to have not survived on the right-hand side of the birth weight cut-off due to lack of policy receipt. Imputations are made varying health outcomes (indicated on the horizontal axis), and fertility (indicated in the legend). All estimation details of each RD model follow those described in Table 2.

D A Back-of-the-Envelope Estimation of Policy Costs and Benefits

Results from sections 5.1-5.2 suggest that the marginal impacts of this intensive early life medical care may actually act to *increase* costs levied in the second generation. This is at odds with the short-term marginal returns documented in previous literature pointing to greater survival (Almond et al., 2010), increased educational performance (Bharadwaj et al., 2013), and reduced links to state support programs (Chyn et al., 2021). It is illustrative to consider what these second generation results imply vis-à-vis the policy's cost, and positive short term returns. The importance of considering such life-cycle returns to social policies has been espoused in García et al. (2020), and here we can move beyond the life cycle of the first generation, into second generation outcomes.

To do this, we consider estimates of the initial costs of providing marginal medical care to treated individuals, the estimated present value of first generation benefits of policy receipt, and the estimated present value of second generation costs. We incorporate into these calculations the welfare cost of financing taxation, and consider a money metric to compare across different policy domains. In each case, we are fortunate to have a directly matched control group to the treated group of individuals impacted by the policy. Namely, the counterfactual outcome for children treated at birth are children who narrowly miss out given the birth weight assignment rule. Nevertheless, we note that this is simply a back-of-the envelope activity, principally to allow us to determine how important the negative intergenerational transmission observed here may be compared with first generation benefits. Below we lay out estimated present values for costs and benefits along with any necessary assumptions to arrive to these figures (calculations are summarised in Appendix Table D1). This exercise is conducted from the point of view of a single marginally treated first generation child who is born in a public hospital, starting from the age of 0, with all costs and benefits discounted to the present day (and expressed in terms of real costs in 2022). In closing this sub-section, we note a number of elements which are likely affected by the initial medical care receipt, but which are not feasible to reduce to a money metric, and which are hence omitted from the calculation of policy returns.

The most obvious program cost which must be covered by public funds are the costs of initial medical care. In the US, Almond et al. (2010) estimate that the cost of marginal medical care provided at this cut-off is 9,450 USD. In order to consider the costs in the context of Chile, we calculate the full costs associated with marginal estimated changes in hospitalisation. The additional use of hospitalisation days is documented in Figure 4a, at approximately 3.5 days in the first year of life, 1.5 days in year 2, 0.8 in year 3 and 0.5 in year 4 (after which no marginal changes in hospitalisation days are observed). Figures estimated from Chile related to the *total* cost of intensive care days (including medical inputs and care) suggest a value of 480,047 CLP in 2011 (Alvear et al., 2013), or approximately 694,000 CLP in current terms, which is equivalent to 855 USD per day.[22] Taking the net present value of costs to the health system at the time of birth, discounted at a 5% discount rate, this gives expected costs of 4,955 USD over the first years of the child's life. If we additionally incorporate the welfare cost of taxation noting the deadweight loss associated with the collection of tax revenues (Feldstein, 1999), suggested by García et al. (2020) as 50 cents on the dollar, this suggests an initial policy cost of 7,433 USD.

Early medical life receipt has been documented to be associated with a number of benefits. We consider here a reduced metric of first generation labour market outcomes, mapping from educational benefits documented by Bharadwaj et al. (2013). Based on their estimated impacts of policy receipt on education in the first generation, and a back of the envelope calculation of the labour market returns to education, they suggest that marginal receipt of intensive health investments at birth may increase incomes by 2.7%. If we discount

[22] Note that Almond et al. (2010) estimate a cost of marginal treatment at birth of 9,450 USD in hospital costs in 2011 USD. Here, based on our estimate of 3.6 additional days of hospitalisation at birth, the equivalent instantaneous cost in USD of hospitalisation in Chile is $855 \times 3.6 = 3078$. Even when incorporating the welfare cost of taxation suggested by García et al. (2020), this value corresponds to 4,617 in 2022 USD, or 3,661 in 2011 USD. Thus costs in Chile are around a third of those in the US, in line with relative cost and hospitalisation usage indexes in the two countries (Lorenzoni and Koechlin, 2017).

the expected flow of future earnings back to birth, based on the median income in Chile in 2021, this suggests directly attributable changes in salaries of 980 USD. Thus, labour market concerns alone do not cover the initial outlay, but if one additionally accounts for the value of a statistical life (VSL), and the marginal likelihood that individuals die during their first year of life, this suggests substantial additional benefits of 142,820 USD (Mardones and Riquelme, 2018).

In addition, we consider the associated costs to the reduction in health at birth of the second generation. We observed three main cost sources: an increase in the probability of receiving treatment in the second generation, an increase in infant mortality, and a decrease in birth weight. We calculate the net present value of the costs and benefits at the time of the mother's birth considering 0.26 additional births up to 26 years documented in Figure 4 for those mothers who were initially treated and the welfare cost of financing taxation. Taking the increase of 0.045 in the probability of being born at less than 1,500 grams and therefore of receiving the treatment, the net present value of the initial medical care costs is 25 USD. Due to the increase of 0.007 observed in the infant mortality rate for those born to treated mothers, a cost of 1,894 USD associated with the value of statistical lives lost must also be included. Finally, we consider changes in future income through changes in schooling. Educational benefits documented by Behrman and Rosenzweig (2004) indicate that increasing the birth weight by 1 lb. increases adult earnings by more than 7%. Thus, the 149 gram decrease in birth weight is associated with an additional cost of 61 USD. The above suggests a second generation cost of the policy of 1,980 USD. In contrast, we observe a decrease in hospitalizations during the first years of life which corresponds to an expected second generation benefit of 155 USD.

These calculations suggest that the total estimated cost of the policy amounts to 9,413 USD while the estimated benefit of the policy is 143,955 USD. If we only consider the 980 USD corresponding to labour returns to education received by the first generation and the 155 USD due to the decrease in early hospitalization for the second generation, the amount is not enough to compensate the initial costs of medical care and the reduction in health of the second generation. However, when incorporating the marginal willingness to pay for the reduction of infant mortality through the VSL, the benefits are around 15 times the total costs. It is important to note that while we can provide a back-of-the-envelope estimate of the policy's net present value, this is necessarily based on a partial picture of the policy's full set of benefits. While we can feasibly value the policy's impact on health, education and future labour market outcomes, there are a number of benefits which we cannot easily quantify. This includes factors such as reduced rates of bereavement given lower rates of infant mortality and potential lower rates of pregnancy loss in the second generation,[23] as well as the policy's impact on allowing women and families greater autonomy to achieve their desired fertility.

It is important to note that due to a lack of recorded father's identifier, our estimated second generation costs and benefits are all based on mother–child links. In general, there is evidence to suggest significant intergenerational links between fathers and their children (Giuntella et al., 2019, for example). The calculations here all refer to costs and benefits for a single individual, and so are not affected if impacts are identical for fathers and mothers. The literature suggests that links between fathers and their children are generally smaller than those between mothers and their children, which, if anything, suggests that second generation costs and benefits would be lower for fathers. Given that costs exceed benefits in the second generation, all told it seems that the inclusion of fathers would only increase the total present value of policy receipt.

[23] Both of these events have obvious and considerable costs to well-being over myriad dimensions (Rogers et al., 2008; Ogwulu et al., 2015; Persson and Rossin-Slater, 2018). However adequately capturing their true value would be difficult.

Table D1: Present value of costs and benefits of policy

	USD	Source
Panel A: First Generation		
Costs		
Medical care per day: 855 USD		Alvear et al. (2013)
Year 1: 3.5 days	2,992.5	
Year 2: 1.5 days	1,282.5	
Year 3: 0.8 days	684.0	
Year 4: 0.5 days	427.5	
PV at birth	4,955.8	
Welfare cost of taxation	2,477.9	García et al. (2020)
Total	**7433.8**	
Benefits		
Labour market returns to education	979.7	Bharadwaj et al. (2013)
Value of a statistical life	142,820	Mardones and Riquelme (2018)
Total	**143,799.7**	
Panel B: Second Generation		
Costs		
Increased probability of VLW	24.5	
Labour market returns to education	60.8	Behrman and Rosenzweig (2004)
Increased infant mortality	1,893.9	
Total	**1,979.2**	
Benefits		
Medical care per day: 855 USD		Alvear et al. (2013)
Year 2: 0.2 days	171.0	
Year 3: 0.5 days	427.5	
Year 4: 0.6 days	513.0	
Year 5: 0.7 days	598.5	
PV at birth	1,415.4	
PV at mother's birth	103.5	
Welfare cost of taxation	51.7	García et al. (2020)
Total	**155.2**	

Notes: Each panel displays the costs and benefits in terms of real costs in 2022. Panel A displays the details for generation 1 and panel B for generation 2 considering 0.26 additional births up to 26 years. All costs are presented discounted to the time at birth of the first generation mother. In the case of benefits, these are reported in terms of time accrued, and then discounted to the time of birth of the mother in 'PV at mother's birth'.