DISCUSSION PAPER SERIES

DP17257

Infant health, cognitive performance and earnings: Evidence from inception of the welfare state in Sweden

Sonia Bhalotra, Martin Karlsson, Therese Nilsson and Nina Schwarz

ECONOMIC HISTORY

LABOUR ECONOMICS

ISSN 0265-8003

Infant health, cognitive performance and earnings: Evidence from inception of the welfare state in Sweden

Sonia Bhalotra, Martin Karlsson, Therese Nilsson and Nina Schwarz

Discussion Paper DP17257
Published 26 April 2022
Submitted 24 April 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Infant health, cognitive performance and earnings: Evidence from inception of the welfare state in Sweden

Abstract

We identify earnings impacts of exposure to an infant health intervention in Sweden, using individual linked administrative data to trace potential mechanisms. Leveraging quasi-random variation in eligibility, we estimate that exposure was associated with higher test scores in primary school for boys and girls. However only girls were more likely to score in the top quintile. Subsequent gains, in secondary schooling, employment, and earnings, are restricted to girls. We show that the differential gains for women accrued from both skills and opportunities.

JEL Classification: I15, I18, H41

Keywords: Infant health, early life interventions, cognitive skills, Education, Earnings, Occupational choice, Programme evaluation, Sweden, gender

Sonia Bhalotra - sonia.bhalotra@warwick.ac.uk
University of Warwick and CEPR

Martin Karlsson - martin.karlsson@uni-due.de
CINCH, University of Duisburg-Essen

Therese Nilsson - Therese.Nilsson@nek.lu.se
University of Lund

Nina Schwarz - nina.schwarz@uni-due.de
University Duissberg-Essen

Infant Health, Cognitive Performance and Earnings:

Evidence from Inception of the Welfare State in Sweden

Sonia Bhalotra, Martin Karlsson, Therese Nilsson and Nina Schwarz

We identify earnings impacts of exposure to an infant health intervention in Sweden, using individual linked administrative data to trace potential mechanisms. Leveraging quasi-random variation in eligibility, we estimate that exposure was associated with higher test scores in primary school for boys and girls. However only girls were more likely to score in the top quintile. Subsequent gains, in secondary schooling, employment, and earnings, are restricted to girls. We show that the differential gains for women accrued from both skills and opportunities.

Keywords: Infant health, early life interventions, cognitive skills, education, earnings, occupational choice, programme evaluation, Sweden, gender

JEL classification: I15, I18, H41

Affiliations and acknowledgements: *Bhalotra:* University of Warwick and CAGE, *Karlsson:* CINCH, University of Duisburg-Essen, *Nilsson:* Lund University and Research Institute of Industrial Economics, *Schwarz:* University of Duisburg-Essen. We acknowledge generous support from The Swedish Foundation for Humanities and Social Sciences (P12-0480:1) and the Swedish Research Council (dnr 2019-03553). We acknowledge partial funding from the University of Warwick and CAGE and partial funding from ESRC Grants ES/L009153/1 and ES/S012486/1. Johanna Ringkvist, Josefin Kilman and Ines Hußmann provided excellent research assistance. We are grateful to Andreas Bergh, Gustav Kjellsson, Alessandro Martinello, Teresa Molina, Cheti Nicoletti, Anton Nilsson, Martin Nordin, Owen O'Donnell, Erik Plug, Peter Sandholt Jensen, Bernhard Schmidpeter and participants of various conferences and seminars for their feedback on earlier versions of this paper. This paper is forthcoming *Review of Economics and Statistics*.

1 Introduction

Spurred by cessation of infant mortality decline, the Swedish government trialled a postnatal intervention from 1 October 1931 to 30 June 1933. It had universal coverage and provided information, support and monitoring of newborn health, including encouragement of breastfeeding, sanitation, a healthy diet and home visiting and clinic attendance. It was a significant pillar in the emergence of the welfare state in Scandinavia. The explicit purpose of the intervention was to bring down infant mortality, important in itself and also as a marker for improvements in infant health for those who survive infancy (Bozzoli et al., 2009). In previous work, we establish that the intervention achieved this goal (a 24% decline in infant mortality) and, in addition, led to meaningful reductions in adult chronic disease mortality (Bhalotra et al., 2017).

In this paper we examine dynamic impacts of the improvement in infant health on educational and economic outcomes. Infancy is a period of rapid neurological development – the brain doubles in size in the first year, and by age three it has reached 80% of its adult volume (Nowakowski, 2006). Brain growth is sensitive to nutrition and infection. It is estimated that 85% of calorie intake in infancy is used to build brains, and severe or repeated infections in infancy may divert nutrients away from brain development (Finch and Crimmins, 2004; Eppig et al., 2010). Moreover, the release of inflammatory molecules during infections may directly impact the developing brain by changing the expression of genes involved in the development of neurons and the connections between them (Deverman and Patterson, 2009). Thus there are biological mechanisms for causal effects of infant health on cognition.

The biological mechanisms may be reinforced as follows. Individuals carrying an improved cognitive endowment from infancy may make greater investments in education, receive reinforcing investments from parents (Yi et al., 2015; Almond et al., 2017; Bhalotra and Venkataramani, 2013; Adhvaryu and Nyshadham, 2016), and compete more effectively for state investments in education (which, we will argue, played a role in the context we study). If the intervention-eligible cohorts exhibit higher human capital attainment and, if there is sufficient demand for the acquired skills, we may expect them to have higher earnings. We will show that, for our sample cohorts, both skill acquisition and opportunities mattered for realisation of the impact of infant health on adult labour market outcomes. This is our main contribution. As highlighted in a survey on the long arm of childhood exposures, the evidence on mechanisms or key levers at different points of the life course is particularly scarce (Almond et al., 2017).

We use linked administrative data for a large and representative sample of individuals tracked from birth, through school, to labour market outcomes and then to retirement and death. It is unusual to have individual longitudinal data from birth to death for a population and, especially unusual to have school test scores linked backwards to quasi-experimental variation in birth conditions, as well as forwards to labour market outcomes. We digitised the birth and school records and linked them to available administrative data on later life outcomes. Individual birth certificate data were obtained from historical parish records for 114 rural parishes and

4 cities that we show were representative of the country in 1930. The population of births in these regions is about 25,000 occurring in 1930–1934. Linkage was done using first name, last name, exact birth date and parish or city of birth. The match rate of births was 66% to school records, 86% to 1970 census files and 65% (91% of survivors) to tax registers. Sample attrition is potentially endogenous because the intervention influenced survival rates but we find it is not differential by treatment status in the census and tax register samples, though it is in the school sample. We nevertheless investigate robustness of all estimates to adjusting for attrition. Match rates were similar for men and women, and attrition adjustments are by gender.

Identification exploits eligibility criteria within treated parishes, and we additionally include matched controls. In treated parishes, children aged 0–12 months at any time in the window for which the programme was available were eligible, for durations that varied with their exact date of birth. We conduct a suite of robustness checks. These include using alternative measures of intervention exposure, investigation of pre-trends, randomisation inference, and sensitivity of the estimates to selective survival. Since there is no variation in treatment timing, concerns about bias due to staggered implementation (cf. Goodman-Bacon, 2018; de Chaisemartin and d'Haultfoeuille, 2019) do not apply in our case.

Our main findings are as follows (magnitudes are reported for a year of exposure). Primary school test scores at age 10 improved for intervention-eligible boys and girls, although with a markedly different distribution of gains. Treatment effects for boys were similar across the distribution (averaging 0.1 standard deviations), while treatment effects for girls were evident only in the upper reaches of the distribution. The intervention increased the chances that girls score in the top GPA quintile by 12.4% points in contrast to an imprecisely determined 2.75% points for boys. These distributional effects were important because secondary school places were limited and the chances of attaining secondary schooling increased sharply for children scoring towards the top of the primary school test score distribution. While primary school attendance was universal by mandate, only about a fifth of all children progressed into secondary school. In line with the intervention shifting girls into the top GPA quintile, we find that intervention exposure is associated with a significant 3.5% point increase in secondary schooling for girls, alongside no change for boys. To illustrate the role of capacity constraints, we leverage arbitrary variation across secondary school catchment areas in the share of treated children, showing crowd-out that disproportionately hurts boys.

Tracking the intervention and control cohorts over time, we observe an intervention-led divergence between the labour market outcomes of men and women at age 36-40. A year of exposure is associated with an average increase in earnings in the treated population of 7.3%, entirely driven by a 19.5% increase in earnings for women. This is large because it reflects an extensive margin increase as well as endogenous increases in skill. As a result of the intervention, women were 5.3% points more likely to participate, they were 7.6% points (20.5%) more likely to be in full-time employment, and almost entirely in the public sector.

Estimates of unconditional quantile treatment effects following Firpo et al. (2009) suggest no gains anywhere in the distribution for men, and that income gains for women were concentrated in the upper part of the distribution. The probability that women belong to the top earnings quintile increased by 8% points. To illuminate this further, we analyse occupational sorting by gender. This reveals that the intervention led to women being 5% points (29.4%) more likely to work as managers and professionals, and 4.4% points (35.5%) more likely to work in accounting, banking and administration. These were among the highest wage occupations in this era. We document the skill intensity of these occupations using three measures.

Using a shift-share approach (Goldsmith-Pinkham et al., 2018; Borusyak et al., 2018), we show significant heterogeneity in intervention effects by an index of the demand for skilled women. In the years in which our sample cohorts were making decisions about higher education and employment, Sweden was experiencing a rapid expansion of the welfare state which created a disproportionate increase in labour demand in public sector occupations dominated by women. We show that health and education were among the top sectors driving the demand for women, accounting for at least 21% of demand growth. We also describe the gender-segmentation of the labour market at this time. We estimate that a one standard deviation increase in an index of demand for skilled women's labour at the parish-cohort level almost doubles the estimated impact of the infant intervention on labour market outcomes for women. This heteroegeneity for women and the average results for men suggest that where growth in opportunities was stunted, treatment effects on labour market outcomes were blunted.[1] We also investigate the role of gender norms and child care expansion, but find no heterogeneity in intervention effects by these variables.

Identifying mediators in longitudinal studies of early life interventions is challenging, requiring either additional sources of exogenous variation or strong assumptions regarding the relationship between variables (cf. Heckman et al., 2013; Heckman and Pinto, 2015; Huber et al., 2017; Dippel et al., 2017). The sequential ignorability condition assumed in some studies is not tenable when, as is common, most outcomes are proxies for human capital. For this reason, it is customary to report the effects of an intervention on potential mediators alongside effects on the final outcomes of interest, without attempting to weight the contributions of alternative mediators. We develop an approach to gauge the extent to which the same individuals contribute to the treatment effects on different outcomes. We complement this with an attribution of treatment effects across potential mediators (Gelbach, 2016). Without claiming to estimate causal mediator effects, we improve upon the standard descriptive representation by using individual data and exploiting the natural sequencing of outcomes across the lifecourse.

To summarise, we find compelling evidence that an increase in primary school test scores lifted up the potential trajectory among individuals exposed to the infant health treatment. Improved labour market outcomes were realised primarily for individuals scoring towards the top of the distribution and, as a result, continuing to secondary school. The results for boys show

[1] 92% of men, 37% of women worked full time. There was limited room for men to increase employment, but they could enter more skilled occupations and increase their earnings.

that capacity constraints can hamper realisation of the full potential of the infant health gain. They also highlight that, under competition, intervention effects on the distribution can determine the size of economic gains. By identifying the skill content of the occupations that eligible individuals entered, we trace a path from infant health to earnings via skill acquisition. We also demonstrate that employment and earnings returns to the intervention varied significantly with growth in the demand for labour. The consistent differentiation of results by gender lends credence to the notion of a causal chain running from earlier to later life outcomes: the higher primary school test scores of women propelled them into higher education, and the growth of skilled occupations in the public sector facilitated the absorption of these more skilled women into the labour market. We provide a crude cost-benefit analysis which suggests a very high internal rate of return to the intervention despite the absent results for men.

Our results are relevant to policy design today. Sweden in the 1930s had an infectious disease environment similar to that in many developing countries today. Modern medicine has progressed but there remains considerable scope to improve preventive measures, including provision of information concerning diet and hygiene and routine checks. While universal health coverage, especially for maternal and child health, is high on the current global health agenda (Gorna et al., 2015), there are few systematic evaluations of immediate or long run impacts (Engle et al., 2007). Early childhood programmes similar to the Swedish trial are being introduced in developing countries, e.g., the Chilean Crece Contigo Programme (Clarke et al., 2018) and the Indian ICDS Programme (Dhamija and Gitanjali, 2019), and being refurbished in richer countries, e.g., the Nurse Family Partnership in the UK (Cattan et al., 2019).

Our findings suggest that a simple low-cost infant health intervention can produce benefits over and above its target, and across domains including infant health, education and earnings. Thus, the return to investing in infant health is much higher than is commonly recognized in global health debates. By virtue of showing that an infant health intervention can improve cognitive skills of children, our results also provide new evidence relevant to what is referred to as a global learning crisis, with millions of children failing to attain their cognitive potential (UNESCO, 2014). Knowledge that differences in cognitive skills emerge early and widen with age has led to a call for pre-school interventions (Flavio and Heckman, 2007; Attanasio, 2015). Our findings suggest an alternative tool with a similar or larger benefit-cost ratio.

Previous work has discussed changes in the relative demand for female (vs male) labour stemming from recession, war, or technological change (Elsby et al., 2010; Acemoglu et al., 2004; Cortes et al., 2018; Bhalotra et al., 2018). We provide a new perspective, emphasising that expansion of broad-based public services tends to raise the relative demand for female labour. This is of potential relevance to understanding prospects for women in developing countries that are witnessing large-scale expansion in the provision of schooling, public health services and pre-school centres. Our findings highlight that the earnings payoff to cognitive skills is uncertain, being dependent upon capacity constraints and demand conditions.

2 Background

2.1 The Field Trial – Institutional Details

The intervention emerged as a potential solution to cessation of maternal and infant mortality decline during the 1920s. It was designed as a trial to precede a decision on nationwide adoption. It ran 1 October 1931 to 30 June 1933 in 7 health districts containing 59 municipalities (2 cities and 57 parishes), chosen to be representative of the country, and not selected based upon infant or maternal mortality rates (see Appendix **??** for details). It was funded by the central government to the tune of SEK 41,400 (USD 139,000 in current prices) (SOU, 1935).

The intervention focused on preventive care and included check-ups at surgeries, home visits and information campaigns. Outreach included announcements in local newspapers and churches, and dissemination by midwives and nurses (Stenhoff, 1934). Every participating district opened a health centre with regular office hours 2–3 times per week. We digitised records maintained by health professionals that indicate programme utilisation. These show that the average infant made 2.8 visits to a health centre and received 3.9 home visits. Newborn children were weighed and checked, and sick children were referred to doctors. Mothers were encouraged to breastfeed and given leaflets illustrating the nutritional needs of children at different stages of development. Home visits by nurses were designed to provide advice on hygiene, sanitation and cleanliness in the household, and to ensure that families followed guidelines published by the National Board of Health. So as to understand what the control group in our analysis received, it is relevant to note that while Sweden had a fairly developed primary care system in 1930, there were limited preventive care and support activities targeting infants and expecting mothers. Annual audit reports in the early 1930s indicate programme fidelity, and harmonisation of activities across the treated districts.

Eligibility for the infant care programme was determined by birth date. All children less than 12 months of age at the start of the intervention were eligible and eligibility ceased on their first birthday. Appendix Figure **??** shows the duration of eligibility by birth date. An antenatal care programme was introduced simultaneously with the postnatal programme and all expectant mothers were eligible. The raw correlation of eligibility duration for the antenatal and postnatal interventions is 0.32 and, conditional on eligibility for any one, it falls to 0.13. Given differential exposure of each individual to the antenatal vs the postnatal programmes, we consistently estimate impacts of the postnatal programme conditional on the antenatal one. In Bhalotra et al. (2017), we found no impacts of the antenatal programme on infant mortality (or later-life health) and here we again find no positive impacts on economic outcomes, so the discussion focuses upon the postnatal programme.

The first systematic evaluation of the trial is in Bhalotra et al. (2017), where we show that the average duration of programme exposure in infancy led to a 1.56% point decline in the risk of infant death (24% of baseline risk) and a 2.56% point decline in the risk of dying by age 75 (7.0% of baseline risk). The intervention-led declines in the risk of dying after the age of 50 were dominated by reductions in mortality from cancer, cardiovascular disease and infections.

2.2 The Swedish School System

In the 1930s, schooling in Sweden started in the year an individual turned seven and was compulsory for six years. Primary education (*Folkskolan*) was universal. Sweden had a tracking system whereby students progressing to secondary schools left *Folkskolan* after either grade 4 or 6. On average, barely 20% of children attended secondary school. A series of reforms 1925–1945, driven by demand and a political will to reduce educational inequalities, increased access and the geographical spread of secondary schools (Lindgren et al., 2014). A 1927 reform granted equal access for girls to all state-led grammar schools and mandated an equal curriculum. The share of girls and boys attending secondary school was similar in 1930. Teachers kept records of test scores and attendance, which we digitised. Govenment guidelines dictated that teachers reward the quality of knowledge and not the quantity, and that grading reflects performance through the year. Teachers were instructed to allow for mark inflation as pupils progressed to higher grades, and to make no adjustment for school form. Thus, the marks reflect an absolute standard and not the relative position of a pupil in their class. For our sample cohorts, schooling was fairly comparable across the country and the curriculum did not change between 1919 and 1950. Appendix **??** provides further details.

3 Data and Empirical Strategy

3.1 Administrative Data Linkage

The data are unique in linking individual-level information across the life course using birth registers, school registers, the 1970 census, and official tax registers. The birth and school registers were digitised by the authors.

Birth Registers. A census of all 24,390 live births in 1930–1934 in the treated and control parishes was digitised from church records, to include births before, during and after the trial of 1931-1933. Sweden is one of the few countries with high-quality vital statistics at the parish level from the 18$^{\text{th}}$ century onwards (Pettersson-Lidbom, 2015). The birth data contain sex, marital status of the mother, age of the mother and parental occupational status (Leeuwen et al., 2002). We merged birth registers with other administrative data using linking procedures described in Bhalotra et al. (2017) in detail.

School Records. We got standardised exam catalogues (Appendix Figure **??**) from regional archives, with yearly pupil-level test scores (math, writing, reading and speaking, religion) and absence in primary school for birth cohorts 1930–34 in grades 1 and 4 (years 1937–47). Grade 1 represents the first possible observation and grade 4 the last as some pupils proceed to secondary schooling afterwards. For about half of our sample we have data for grades 1 and 4, for the rest one or the other. We also have length of school year and school type.

Birth register data were matched to school records using parish, birth date, and full name. Of 22,500 individuals alive at age 7, roughly 16,000 were matched. Missing information is mostly because archives of certain schools were accidentally destroyed. We mitigated data lost on account of migration by tracking migrants and collecting school records from destination parishes, even when these were not within our main sample of treated and matched parishes.

Labour Market Outcomes. We matched individuals in the birth records to the 1970 census which covers the entire population of Sweden on 1st November 1970 (SCB, 1972). It contains educational attainment, income, employment status and occupation. Of 24,390 births, we match roughly 20,900. Of the 3,490 unmatched individuals, we discerned by a matching of births to death registers that 74% died before 1970.[2]

Pension Income. The match to pension income for 2001–2005 from official tax registers succeeds for 16,180 individuals. Of the unmatched, 6,621 died before 2002, leaving 1,589 survivors unmatched. We use pension income because it is insensitive to career interruptions such as around childbearing, which could influence 1970 income given that the intervention cohorts were of reproductive age then. In Sweden at this time, a full pension required thirty years of contributions and the amount was based upon the best fifteen years (Sundén, 2006).

Longitudinal Individual Data: Four Points in the Lifecycle. To summarise, we track outcomes at four points in the lifecycle. The potentially treated cohorts are born 1931–1933, and observed in grade 1 at age 7, in grade 4 at age 10, with labour market outcomes age 37–39 and pensions age 71-73. Conditional on survival, the match to birth data is 72% for school registers, 96% for labour market outcomes and 91% for pensions. Match rates are similar for men and women. Appendix Table **??** provides attrition rates by intervention eligibility. Appendix Tables **??** and **??** present descriptive statistics.

Matched Controls. We identified as matched controls, 2 cities and 57 rural parishes (belonging to 38 different health districts) using observable parish characteristics from the 1930 census and the Mahalanobis distance metric; see Appendix **??**, where we also present validating tests and demonstrate that our analysis sample is representative of Sweden. The matching is also detailed in Bhalotra et al. (2017).

3.2 Empirical Strategy

For identification, we leverage a discontinuity in eligibility criteria. Children were eligible between birth and the age of 12 months, which delivers variation in eligibility *within* treated parishes. We additionally use matched controls, so that the estimates are derived from comparing outcomes for exposed cohorts in treated regions to those of unexposed cohorts and control regions. In contrast to the case in most DID designs, the intervention is switched on and off, as a result of which unexposed cohorts include ineligible individuals born before and after the ex-

[2]Earnings data in the 1970 census is regarded to be of high quality, but women who were partners of a small business owner could be recorded as working while having zero taxable earnings. We impute incomes of these 2,987 women based on qualifications and hours worked.

posed cohorts. We define exposure as duration, using exact date of birth together with the exact dates of the start and end of the intervention. In robustness checks we investigate alternative formulations including one that accounts for age of initial exposure. The estimated equation is:

$$y_{ipt} = \alpha + \beta T_t + \gamma_p + \tau T_t D_p + \sigma_t + \lambda X + u_{ipt} \tag{1}$$

where y_{ipt} is the outcome for child i born in parish p on day t, T_t is the duration of eligibility (in years) for the intervention for birth date t, D_p is a dummy equal to one for treated parishes, γ_p are parish fixed effects, σ_t are $Quarter\ of\ birth \times Year\ of\ birth$ fixed effects and X is a vector of covariates. Covariates include fixed effects for a hospital birth, mother's marital status, twin birth, older (>35 years) and younger (<25) mothers, and the occupational status of the household head at the birth of the child. We control for eligibility for the maternal intervention and for two school reforms increasing instructional time (see Appendix **??**). When modelling school outcomes we control for school fixed effects, length of school year, and school form, an indicator of school quality. Since eligibility is determined by the same function of birthdate for all treated children, our estimates are unlikely to be confounded by secular trends; besides, they are immune to the problems that arise when DID is applied to staggered implementation (Goodman-Bacon, 2018; de Chaisemartin and d'Haultfoeuille, 2019). We nevertheless investigate robustness to including parish specific time-trends, and to including health district fixed effects and health district specific trends. We scrutinize pre-trends in event study plots.[3]

The parameter τ captures the intent-to-treat (ITT) effect of the infant intervention for an additional year of eligibility. This is the parameter of interest for policy makers who are unable or unwilling to make the utilisation of services mandatory. Since there were no always-takers (cf. De Chaisemartin, 2012) the ITT is a scaled version of the average treatment effect on the treated (ATT). As in all studies of the long run effects of a positive health intervention, surviving individuals are negatively selected and, as a result, our estimates will be conservative.

In Bhalotra et al. (2017) we presented results which increase our confidence that programme variation across birth cohort and birth parish is quasi-experimental. We showed that school reforms in the study period (see Appendix **??** and Fischer et al., 2019) are orthogonal to the infant intervention, but we nevertheless control for them. We also showed that programme impacts on infant and adult health were not sensitive to accounting for trends in hospital birth or parish-specific impacts of the Great Depression. Furthermore, Sweden was neutral in World War II and historical sources suggest no educational disruptions for our sample cohorts (Fredriksson et al., 1971), and we confirm no structural breaks in our school data in the war years.

Section 7 and Appendix **??** present a number of specification checks for the long run outcomes that we analyse in this paper. We examine alternative definitions of the treatment indi-

[3]Goiter, caused by iodine deficits, has irreversible effects on brain development and was common in Sweden in the 1930s. Iodine fortified salt was introduced in 1936, after the end of our intervention. Inclusion of parish fixed effects and trends accounts for baseline geographical variation in the level of iodine, and possibly differential impacts of iodized salt on cognitive development across parishes with different initial levels.

cator, sensitivity to covariates, anchoring of school grades to adult income, balance on baseline covariates and tests for differential pre-trends. We also identify and allow for differential attrition by treatment status, and implement placebo and randomisation inference tests.

4 Results

In this section we start by creating indices of outcomes at the three stages of life and adjusting for multiple hypothesis testing. We then present the main results for primary school test scores, progression to secondary school (higher education in this era), and earnings. This is followed by analysis of sickness absence in school, and employment and occupation in adulthood as proximate sources of changes in education and earnings.

4.1 Outcome Indices

The outcomes we analyse fall into a life course hierarchy with earnings (or pension) being the primary endpoint. Nevertheless, the many estimates we present might potentially raise concerns about false discoveries. To safeguard against this, we subject results for multiple outcomes measured at a given age to a multiple hypothesis testing correction. Since many of these outcomes are strongly correlated for substantive reasons (e.g. GPA determines secondary schooling entry, and occupation is predictive of earnings), we follow Anderson (2008) and construct indices that take these correlations into account: the Age 7 Index includes GPA and top quintile GPA in grade 1, the Age 10 Index includes top GPA in grade 4 and secondary schooling and the Adult Index includes top income, log income, log pensions, full time employment, municipal employment, federal employment, and an indicator for high-ranking occupation.[4] Results based on our main specification are in Table 1.

We find a statistically significant improvement in educational outcomes at age 10 of 0.13 standard deviations (SD) and in adult labour market outcomes of 0.07 SD for women, but not men. For each estimate, we present conventional p-values alongside p-values from a multiple testing adjustment controlling the false discovery rate. For this we follow the two-stage procedure of Benjamini et al. (2006). The adjusted p-values are higher but the results remain significant at the same significance levels as indicated by conventional standard errors.[5] Figures 1(a)-1(b) present event studies for females for the age 10 and adult indices that were affected by the intervention. These plots confirm the main results and lend support to the empirical strategy. Similar plots for the component outcomes, discussed below, are cleaner.

[4] Top GPA and top income are indicators for being in the upper 20% of the distribution. Section 4.2.1 suggests gains in mean GPA favour boys while gains in top GPA favour girls. Since we consistently find significant outcome differences between the genders, we use top GPA rather than GPA in the Age 10 Index, the age at which we find treatment effects. If we additionally include GPA, results are noisier but the broad patterns are similar.

[5] We conduct the adjustment for all six tests jointly, whereas Anderson (2008) does it by gender. Our adjustment is thus more conservative but since only two parameters are significant using conventional p-values, splitting by gender would not change results.

4.2 Outcomes: Human Capital and Earnings

4.2.1 Cognitive Performance - Primary School

We created a measure of cognitive ability by taking the mean of grades in math, reading, and speaking and writing to form a grade-point average (GPA), though we also report subject-specific estimates. Girls got better marks than boys, and marks in grade 4 exhibit a higher mean and greater spread than in grade 1. To ease interpretation of the coefficients we transform marks into a z score using the inverse standard normal distribution. Appendix Figure **??** plots these data by gender and grade.

Quantile treatment effects. Figures 2(a)-2(b) plot unconditional quantile treatment effects for grade 4 GPA by gender. The corresponding plots for grade 1 GPA showed no significant treatment effects, in line with Table 1. Following Firpo et al. (2009), quantiles are defined pre-regression and covariates help adjust for selection bias without redefining the quantiles (see e.g. Borgen, 2016; Killewald and Bearak, 2014). Boys experienced positive treatment effects across the distribution. For girls, the upper 30% of the distribution is significantly higher on account of the treatment. The gender difference in impacts of the treatment on the test score distribution translates into a gender difference in access to secondary school – see Appendix Figure **??** which shows that secondary school completion rates increased sharply with primary school test scores towards the top of the score distribution. Since, on average, 20% of the sample cohorts attained secondary schooling, we obtain regression estimates at the mean and for the probability of scoring in the top quintile of the pooled (male and female) GPA distribution.

Mean and distribution of test score gains. A year of exposure to the intervention leads to a statistically significant increase of about 0.08 SD in average GPA in the full sample. Table 2 shows results by gender. The coefficients are not significantly different by sex, but larger and only statistically significant among boys, who exhibit a GPA increase of about 0.11 SD. The intervention increases the probability of being in the top quintile by 7.5% points on average, and by 12.4% points for girls in contrast to an imprecisely determined 2.75% points for boys.

Subject-specific estimates show that average GPA improvements are in 'writing' and 'reading and speaking', which increase by about 0.11 and 0.12 SD. These coefficients are not significantly different by gender but larger and only statistically significant for boys.[6] We find no discernible impact of the programme in grade 1 (Appendix Table **??**). As some recent studies find that cognitive gains stemming from pre-school interventions fade (see e.g. Bitler et al., 2016; Chetty et al., 2011), while theory predicts that the gains will multiply over time, it is notable that the infant health intervention we consider produced cognitive gains that only become

[6]Levine and Schanzenbach (2009) refer to Jacob (2005) to argue that differences in programme impact by subject may arise if performance in some subjects is more sensitive to the value added by school inputs than to other inputs like family environment or initial health. Our finding that effects on literacy dominate effects on math is also seen in e.g. Sievertsen and Wüst (2017) and Aizer et al. (2018), but there are studies showing similar responses of math and reading (e.g. Figlio et al., 2014; Almond et al., 2014). Thus, we are not alone in finding differences but there appears to be no clear scientific explanation for them.

evident at age 10-12. Examining heterogeneity we find that children born out of wedlock benefited substantially more than other children but there were no differences in effects by parental socio-economic status (see Appendix **??**).

Our estimates are sizeable and suggest that pre-school health interventions have the potential to raise cognitive attainment as much as interventions that directly target cognitive capacity. In Appendix **??**, we demonstrate this with reference to effect sizes from related studies.

Event study plot. Figures 1(c)-1(d) present event study style plots for the Top GPA outcome by gender, showing coefficient estimates for each quarter of birth. This confirms that improved school performance coincided with treatment eligibility for girls, whereas no such relationship is discernible for boys.

4.2.2 Secondary Education

The upper panel of Table 3 shows that an additional year of exposure resulted in a 3.5% point (17.6%) increase in the probability that girls completed secondary school, while there was no change among boys. Note that the control group mean is not significantly different between boys and girls. Only 20% of children continued to secondary school and places were rationed. Progression to secondary was a function of primary school test scores. Our result that the intervention led to girls being more likely to score in the upper part of the distribution (Figures 2(a)-2(b)), together with their baseline performance being stronger can explain our finding that only secondary school attainment rates for girls improved. Also, see Appendix Figure **??**.[7]

The greater entry of intervention-eligible girls to secondary school is also consistent with higher returns to secondary school among girls. We show this in Appendix Table **??**.[8] Our finding that an early life health intervention led to better educational outcomes for girls than for boys is consistent with the predictions of Pitt et al. (2012), premised on men having a comparative advantage in brawn-intensive activities. Later we show occupational sorting by gender consistent with this. Bhalotra and Venkataramani (2013) find broadly similar results in Mexico in the 1990s and Saaritsa and Kaihovaara (2016) in Finland in the early 20$^{\text{th}}$ century. Additional explanations of larger treatment effects for girls, that we are unable to test with our data, include that the cognitive growth curve differs by gender, that non-cognitive skills such as conscientiousness that are complementary with cognitive skills enhanced girl effort for a given increment to the cognitive endowment, and that the lower labour force participation rates of women in this era motivated girls to work harder to succeed.

[7]Eligibility for secondary school involved doing well on a national entrance test and passing all subjects in primary school (Appendix **??**). According to Skolöverstyrelsen (1955), about 11% of applicants in our cohorts were rejected. This is relevant to understanding why the intervention did not raise secondary schooling for boys, even though it did improve their performance on average. We examine this carefully below.

[8]Cf. Björklund and Kjellström (1994); Bång (2001). Lifetime returns to education increased for women after a 1939 reform which prohibited firing women on grounds of marriage or pregnancy, similar to the lifting of marriage bars in the United States (Goldin, 1988).

4.2.3 Earnings

We estimate that a year of exposure to the intervention raised earnings by 7.3% on average, driven by women experiencing an increase of about 19.5%, in contrast to no gain among men (middle panel of Table 3). Unconditional quantile treatment effects show no earnings gains anywhere for men but that for women the upper part of the income distribution is moved upward (Figures 2(c)-2(d)), similar to the pattern observed for cognitive performance. The probability of belonging to the top quintile of earners increased by 7% points among women.[9]

The large earnings increase among women is plausible because (a) eligible women acquired stronger skills and (b) it includes an extensive margin increase, which we estimate can account for four-fifths of the observed increase.[10]

Pension income. Since earnings are measured in 1970, when the sample cohorts are 37-39 years old, they may be sensitive to lifecycle variation in labour supply, important for women on account of fertility. Pensions mirror the best fifteen years in the labour market and thus represent earnings at advanced stages of the career. Investigating pension income at age 71 as an alternative measure of earnings, we estimate an increase of 7% for women, and no increase for men (middle panel of Table 3), ratifying the earnings results. We show that this result is robust to controlling for individual receipt of a widow pension, which would create a wedge between women's earnings and pension income (Appendix Table **??**).[11]

Internal rate of return. The intervention cost approximately SEK 41,400 (USD 139,000 in current prices). Personnel costs (salaries for physicians and nurses) accounted for 50% of total costs. The cost per treated child was about USD 39 (in current prices) and per consultation USD 5.7 (in current prices). These costs are low relative to the benefits we identify. We calculate the net present value of earnings and estimate the internal rate of return on the funds spent by the national government as 0.22, see Appendix **??**.

Event study plot. Figures 1(e)-1(f) plot event studies for top income by gender. A greater probability of being in the top income quintile coincides with treatment eligibility for women but not for men, similar to the pattern we identified for treatment effects on top GPA.

---

[9] While cognitive gains apparent for boys did not impact on their earnings, they may have affected domains we do not measure, e.g. financial decision making or marital stability.

[10] Suppose that prior to the intervention, n_2 individuals work full-time, n_1 work part-time and $1 - n_1 - n_2$ do not work. Their log earnings are y_2, y_1 and y_0, respectively. After the intervention, n_2^1 individuals work full-time and n_1^1 work part-time. The extensive margin effect on earnings may then be calculated as

$$\frac{\Delta y}{y_0} = \frac{\left(n_2^1 - n_2\right)\left[\exp\left(y_2\right) - \exp\left(y_0\right)\right] + \left(n_1^1 - n_1\right)\left[\exp\left(y_1\right) - \exp\left(y_0\right)\right]}{n_2 \exp\left(y_2\right) + n_1 \exp\left(y_1\right) + n_0 \exp\left(y_0\right)} \tag{2}$$

In our case, $n_1^1 - n_1 = 0$, $n_2^1 - n_2 = 0.076$. There were large differences in female earnings by employment: $y_2 = 9.89$, $y_1 = 9.18$, $y_0 = 7.93$. Hence, we get: $\frac{\Delta y}{y_0} = \frac{0.076 \cdot 16{,}953}{8{,}022} = \frac{1{,}288}{8{,}022} = 16\%$

[11] For men, we estimate a decline in pension income of 4%. Since we saw no decline in earnings for men at age 37-39, and since only 63% survive to the age of 75, this may reflect endogenous survival selection. To investigate this, we re-estimated programme effects on 1970 income for subsamples of individuals surviving until age 40, 50, 60, 70 and 75 respectively (Appendix Table **??**). We see no selection among females until age 75, when there appears to be some positive selection. Among men, there appears to be negative selection from age 60 onwards as the earnings estimates become progressively lower the older the age group.

4.3 Intermediate Outcomes - Illuminating Mechanisms

4.3.1 Sickness Absence in Primary School

There are two channels through which infant health may have had the observed impacts on school performance. First, infant health may predict school-age health, creating a contemporaneous effect on test scores because healthy children miss school less often or concentrate better when at school. The second channel operates through health facilitating brain development in infancy (Eppig et al., 2010). We investigate intervention effects on sickness absence, a marker of school-age health, with a view to discriminating between the channels. Focusing on grade 4, Appendix Table **??** and Figure **??** show that the intervention reduced sickness absence for boys by about 0.8% (20%), and it is possible this contributed to their higher GPAs. However, we see an unexpected increase in sickness absence for girls.[12] Given that we find increases in secondary school and earnings for girls and not boys, this undermines the relevance of the pathway involving morbidity in the school years in favour of the argument that the intervention improved neurological development, though it is not conclusive.

4.3.2 Employment

Women exposed to the intervention for a year exhibited an increase in the propensity to work full-time of 7.6% points, or 20.5%, and no change in the propensity to work part-time which is 20-35 hours (lower panel of Table 3). More women joined the labour force, this is explicit in the next section. There are no significant impacts on employment for men, 92.5% of whom worked full-time. In the years when our sample cohorts were making the relevant decisions, there was a substantial expansion of the welfare state and a sharp increase in the share of working married women (Stanfors, 2003). We posit that these phenomena are related on account of the growing welfare state creating more jobs for women than for men, as nurses, teachers and child-care workers. We will demonstrate in Section 5.2 that intervention-treated individuals emerging on the market with enhanced skills were more likely to translate their skills into earnings in regions where the demand for skilled workers was expanding. So treated women did better than treated men both because they acquired more skills and because the skilled labour demand increased more in jobs dominated by women.[13] In the next section we estimate treatment effects on occupational sorting, which buttress this argument.

4.3.3 Occupation and Skill

Public Sector Jobs. We find that a year of eligibility for the intervention is associated with an increase in the probability that women work in municipal jobs of 4.9% points, or 20.5%

[12] An improvement in health for boys relative to girls is consistent with the stylized fact of boys being more sensitive to health inputs in infancy, although baseline sickness absence rates are similar for boys and girls.

[13] Cf. Coles and Francesconi (2017) who argue that expanding job opportunities for women was critical for the impact of the contraceptive pill on women's outcomes in America.

relative to the baseline of about 24%, and an increase in the probability of working in federal government jobs of 3.4% points, or 66.5% (lower panel of Table 3). Adding up across these categories, it is clear that all of the additional employment of women was in the public sector. This lines up with our hypothesis that rapid growth of the welfare state created jobs for women (see Appendix Figures **??** and **??**).

Occupation. Examining treatment effects on 2-digit occupations, we find that the increase in women's employment was concentrated in high-skilled sectors (Appendix Table **??**). Women exposed to the intervention for a year were 5.0% points (29.4%) more likely to work as managers and professionals and 4.4% points (35.5%) more likely to work in accounting, banking and administration (almost all of this increase is as office workers or administrators). In contrast, we see a reduction in the share of men in the managers, professionals category and an increase in the share of men in sales. Appendix Table **??** shows the share of men and women in 1970 in each of the main 2-digit occupations, their mean earnings and skill intensity by three independent indicators: average GPA in the occupation, share of workers with secondary education, and the average task content classified as routine vs non-routine cognitive vs non-cognitive following Autor et al. (2003). The two highest-ranked occupational groups (Managers & Professionals and Accounting, administrative) are high-skilled by all three criteria. The highest-paying occupation was managers and professionals – so these results are consistent with our finding that treatment led to women being more likely to appear in the top quintile of the earnings distribution.[14] Disaggregating further, we find the largest increase in the manager/professional category derives from women working in the health sector, as midwifes or nurses (Appendix Table **??**). These results also confirm that the increase in employment of women included an increase in labour force participation, with no similar change for men.

We further investigate 3-digit occupations to identify the highest-earning occupations by gender. We divide the sample into individuals with earnings in the top-20 and the remaining 80% of the distribution. We then rank 3-digit occupations by their employment shares, retaining the eight occupations that employed half of all individuals in the top quintile, see Appendix Table **??**. We see a close correspondence between top quintile earnings and public sector employment for women but not for men. There is also clear evidence that the labour market was gender segregated, the top-ranked occupations for women and men exhibiting limited overlap.

5 Mechanisms – Skill Acquisition and Opportunities

We now extend the analysis to add additional evidence for our argument that the intervention led to increases in higher education and earnings for women but not men, namely, that women experienced (i) greater skill acquisition, and (ii) stronger growth in labour demand. Here we demonstrate that (i) treatment was associated with women displacing men in secondary school

[14]Mining and Crafts shows a higher return for women though not for men. We disregard this aberration as 0.1% of women are in mining and 0.6% in crafts.

admission and (ii) treatment effects were stronger in municipalities in which predicted demand for skilled women was greater. As the intervention had similar impacts on infant mortality of girls and boys (Bhalotra et al., 2017), it is unlikely that the long run differences are a trace of larger (first stage) programme impacts on girls.

5.1 Skill Acquisition

Treated boys had large significant improvements in school performance at age 10, yet their secondary school attendance and labour market outcomes did not improve. Girls were more likely to appear in the upper regions of the test score distribution, and this made them more competitive for rationed secondary schooling places. Here we demonstrate that secondary school capacity constraints disproportionately hurt boys. An alternative explanation may be that treated boys faced higher opportunity costs, but Fischer et al. (2019) show that only a small fraction of our sample cohorts entered regular employment after finishing compulsory schooling.

In the early 1940s, when our subjects were potentially entering secondary schooling, each of the 2,500 municipalities (parish/city) in Sweden had a primary school. However, only 194 had a lower secondary school (Lindgren et al., 2019), and we checked that this provision did not respond to the intervention. The share of each cohort attending secondary schooling was 7 to 42% at the parish level, and most of this variation was between catchment areas. The infant intervention was delivered at the health district level. There were 400 health districts and thus, importantly, fewer secondary school locations than health districts. As a result, the competition faced by a treated child in gaining entry to secondary school varied arbitrarily as a function of the share of treated children in the catchment area of the secondary school closest to them. There is some additional variation from cohort size and birth date.

We exploit this arbitrary variation in order to test the crowding-out mechanism. For each individual i, we measure exposure to other treated children as

$$Exp_i = \frac{\sum_{j:p_j \in CA_{S_{p_i t_i}}} T_{t_j} D_{p_j} 1\left(d_{p_j S_{p_i t_i}} < 50\right)}{\sum_{j:p_j \in CA_{S_{p_i t_i}}} 1\left(d_{p_j S_{p_i t_i}} < 50\right)} \tag{3}$$

where p_i and t_i represent the birth parish and birth date of individual i, $S_{p_i t_i}$ is the nearest secondary school from parish p_i at date t_i, and $CA_{S_{p_i t_i}}$ is the catchment area of that school for birth date t_i. We assign each parish to the catchment area of the closest secondary school, with distance calculated as great-circle distance between parish centroids. T_{t_i} is duration of eligibility and D_{p_I} is a dummy taking on the value 1 if parish p_i was treated. Since most pupils in a secondary school came from within a radius of 50 kilometres, we calculate exposure imposing this restriction. Appendix Figure **??** plots the identifying variation – the share of treated children in the catchment area of the nearest secondary school for a treated child. Appendix Figure **??** confirms that children in the control group were not exposed to treated children.

Figure 3 shows that the treatment effect on secondary school completion varies with the share of treated children. The horizontal axis represents percentiles, n, of the distribution. The coefficients show treatment effects by gender on secondary schooling when the share of treated children is at least n. The figure provides evidence that treatment led to crowd-out and, moreover, that boys suffered crowd-out more than girls. Once the share of treated children crosses 20%, there is a significant decline in the probability that a treated boy completes secondary school. A treated girl also suffers competition, but the decline in probability is smaller and only statistically significant once the share of treated girls exceeds 65%.[15] Appendix Table **??** shows that this rationing of secondary school places is mirrored in male labour market outcomes in the long term. We investigate labour market outcomes from a different angle in Section 5.2, highlighting that while there was possibly some displacement, it was not a zero sum game at the labour market level, albeit more favourably for women.

5.2 Growth in Labour Market Opportunities

Below we show evidence of (i) occupational gender segmentation, (ii) stronger employment growth in sectors dominated by women, (iii) long run intervention effects increasing in parish-level employment growth, but only for women, and (iv) that this result is driven by skilled women, in line with our earlier results showing that treatment raised skills and led to women entering high-skilled occupations.

Occupational gender segregation is clear in Appendix Table **??**. Women were significantly more likely to work in high-wage (skilled) public sector occupations. Three of the eight most populous occupations in the top quintile of the earnings distribution for women were teacher, nurse and medical assistant – which accounted for a third of all women in the eight occupations in the top quintile, in contrast to which only 6% of men were public sector workers (teachers). In this era, it is notable that teachers earnings were in the top quintile of the male earnings distribution, see Figure **??**. The occupations that dominate the top earnings quintile for men are engineers and architects. Understanding sectoral compositional change is clearly relevant to understanding the growth in women's employment at the top of the earnings distribution.

We leverage plausibly exogenous variation across parishes in baseline industrial composition to predict the growth of demand for female and male labour using a Bartik shift-share approach (Goldsmith-Pinkham et al., 2018), Borusyak et al. (2018). The procedure is detailed in Appendix **??**, where we also plot the distribution of the Bartik index. Appendix Table **??** shows that the index is predictive of employment in 1970. We interact the treatment term describing exposure to the infant intervention with the skilled worker Bartik index of own gender, and in columns (2) and (4) we additionally include an interaction with the Bartik index for the opposite gender. The latter allows us to interpret the own-gender interaction as conditional on general labour market conditions. It also tests for spillover effects – for instance, whether

[15] Appendix Table **??** shows estimates collapsed at 50%. Results may be confounded by effect heterogeneity if it is correlated with the share of treated children. But such heterogeneity is unlikely to give rise to the monotonous relationship observed for both genders in the figure.

employment growth in a female-dominated sector affects male labour market outcomes. The Bartik indices are in z scores, so that the estimate for *Treated* $\times$ *Duration Eligibility* represents the treatment effect at the mean of the index.

The results show that effects of the infant intervention are increasing in predicted labour demand for skilled females at the parish-cohort level. On average, this is not the case for men, though we do see positive effects for men born in parishes with very large values of the Bartik index. There is considerable heterogeneity among women. Increasing the Bartik index by 1 SD almost doubles the effect of the intervention on long-run outcomes for women and, conversely, women born in parishes in the lower tail of the index experienced no gains from the intervention. In Appendix Table **??** we show that this same pattern of heterogeneity holds for the range of labour market outcomes, although with varying levels of statistical significance.

The interaction term involving the opposite gender Bartik is not significant – women's labour market outcomes are not sensitive to male employment growth at the parish level, and vice versa. This is consistent with gender segmentation of the labour market. Repeating the analysis using indices for all as opposed to skilled workers, there is no longer any evidence that labour demand is a significant moderator of the effects of the infant intervention (Appendix Table **??**). This is consistent with the intervention having influenced labour market outcomes via skill acquisition. This was a time when there was growth in demand for skilled workers (and, as shown before, this happened to be stronger in woman-dominated occupations in the public sector). Rotemberg weights computed by decomposing the Bartik estimator into a weighted sum of the just-identified instrumental variable estimators, that use each industry share as a separate instrument, reveal that the health care and education sectors were in the top-5 sectors predicting demand for skilled women, accounting for at least 21% of demand growth, see Appendix Table **??**.

Overall, our findings suggest that the average earnings payoff to infant health was uncertain, being dependent upon capacity in higher education and on context-dependent labour demand. In our setting, both constraints acted to limit intervention-led gains for men.

5.2.1 Gender Norms and Childcare Expansion

We have established that growth in opportunities for women was an important mechanism. We now investigate whether changing gender norms facilitated translation of skills into earnings among women. Gender norms are only ever measured by partial indicators, and they tend to evolve slowly. We use baseline differences between parishes in female representation on local councils as a proxy. This is supported by Beaman et al. (2012), who show that gender quotas in local government lead to higher aspirations and educational attainment among girls, demonstrating tangible effects on women's economic participation.

During the first decades of the 20$^{\text{th}}$ century, Sweden made a transition from predominantly direct democracy to representative democracy at the local level (Hinnerich and Pettersson-

Lidbom, 2014). We digitised data (Statistics Sweden, 1947) on the number of elected women and seats in each municipality level council for the 1946 election, the first local election in which a sufficient number of municipalities were practising representative democracy, defining an indicator for above-median female representation in the birth parish. The median share of women in local councils was 0.07, and 20% of our sample were exposed to zero female representatives. We may be concerned that women's political representation in 1946 is an outcome of the intervention in 1931-1933. However, we leverage identifying variation within treated parishes using eligibility by birth date, and then interact this with differences in female representation between treated parishes. The estimates for the adult index are in Appendix Table **??**, and estimates for all labour market outcomes in Appendix Table **??**. There is no evidence that local gender norms, as captured by women's political power, moderated translation of infant health improvements into labour market outcomes for women.

We also investigate the role of childcare in moderating treatment effects on women's labour market outcomes, leveraging an expansion of state-subsidised childcare from 1963. We find no heterogeneity in intervention effects, see Appendix **??**.

6 Mediators – Tracing and Decomposing Effects

Identifying mediators is a central challenge in longitudinal studies of early life interventions (Heckman et al., 2013), requiring either two sources of exogenous variation or strong assumptions regarding the relationship between treatment, mediators and main outcomes. For this reason, it has been customary to report the effects of an intervention on potential mediators alongside effects on the final outcome of interest, without attempting to weight the contributions of alternative mediators. This is what we have done so far. In recent studies,identification has been premised on a sequential ignorability condition, whereby unobservables that confound the relationship between the treatment and the mediator are different from those that confound the relationship between the mediator and the outcome, conditional on treatment (cf. Heckman and Pinto, 2015; Huber et al., 2017; Dippel et al., 2017). This independence assumption is implausible in this setting, where the outcomes are all proxies for human capital. We develop a simple approach to gauge the *relatedness* of the treatment effect over different domains, examining whether it is the same sub-populations that contribute to the treatment effects at different stages of the lifecycle.[16]

To complement this analysis, we leverage the omitted variable bias formula to attribute treatment effects across potential mediators, following Gelbach (2016). The Gelbach approach is essentially agnostic about the causal and temporal ordering of potential mediators. Thus, if the treatment effects on different mediators are strongly correlated, the method may deliver misleading results. We (partially) address this pitfall by using the natural sequencing of outcomes

[16]Deuchert et al. (2016) is similar but requires observing the value of the mediator for treated individuals before treatment, and identification is based on this mediator having no effect on the outcome in the pre-treatment period. Thus, it cannot be applied in our design.

revealed by our analysis of correlated effects to formulate the specification for the Gelbach (2016) decomposition.

6.1 Relatedness of Treatment Effects

In Appendix **??**, we show that the estimated average treatment effect on an interaction between two binary outcomes (i.e. $Y = W \cdot Z$), denoted τ_Y, carries information on how strongly the treatment effects within the domains defined by the two binary outcomes W and Z relate. First, we may compare τ_Y to the benchmark value τ_Y^{uc} that it would take if the treatment effects in the two domains were completely unrelated at the individual level:

$$\tau_Y^{uc} = \tau_W \tau_Z + \tau_W \Pr\left(Z^0 = 1\right) + \tau_Z \Pr\left(W^0 = 1\right), \tag{4}$$

where τ_W and τ_Z are the average treatment effects on the two outcomes W and Z and $\Pr\left(Z^0 = 1\right)$ is the (estimable) counterfactual probability of observing $Z = 0$ in the treatment group in the absence of treatment, and $\Pr\left(W^0 = 1\right)$ is analogously defined.

Table 5 shows the relatedness of the treatment effects of the intervention for outcomes that exhibited significant results for women. The first row shows that intervention exposure is associated with an increase in the probability of scoring a high GPA in primary school (grade 4 top 20%) of 10.55% points, and an increase in the probability of secondary schooling of 5.2% points.[17] The third column presents τ_Y^{uc}, the benchmark value of τ_Y, which is the effect on the interacted outcome (top GPA *and* secondary schooling), which would be obtained if the treatment effects were uncorrelated. In this example, this is 3.4% points. However, the unrestricted treatment effect for this joint outcome, presented in column (4), is almost twice that number, indicating that the treatment effects on the two outcomes are strongly correlated. The last three columns present the correlation coefficients between the two treatment effects.

Table 5 exhibits some striking patterns. First, the estimated value of τ_Y is always above the benchmark value τ_Y^{uc}, typically twice as large, suggesting that treatment effects are strongly correlated for all pairs of outcomes. Treatment effects on earnings are highly correlated with treatment effects on each of high-ranking occupation (0.60), secondary schooling (0.58) and top GPA (0.54). See Appendix **??** for a derivation of the correlation coefficient. The results suggest that the infant intervention benefited the same sub-population in terms of different outcomes across the life course. Since the trajectory of boys and girls diverges at secondary school entry, we next estimate a relatively flexible specification that introduces interactions with secondary schooling.

[17] These are essentially the results from the previous section, slightly different because due to the requirement that both outcomes are observed for a given individual.

6.2 Gelbach Mediation Analysis

We follow Gelbach (2016) to estimate the relative contribution of endogenous outcomes at different stages of the lifecourse to earnings in adulthood. Denoting by Y a $N \times 1$ vector representing top earnings and by T the $N \times 1$ a vector of treatment assignment, we compare results from a specification in which all potential mediators Z are included as covariates with a specification which only includes the base covariates and fixed effects:

$$Y = T\tau + X\lambda + \epsilon \tag{5}$$

$$Y = T\tau + Z\beta + X\lambda + \upsilon \tag{6}$$

Let $\hat{\tau}_{base}$ denote the estimate of τ based on specification (5), and $\hat{\tau}_{full}$ denote the estimate of τ based on specification (6). Their difference $\hat{\delta} = \hat{\tau}_{base} - \hat{\tau}_{full}$ is an estimate of how much of the effect of treatment on earnings can be attributed to the mediating variables Z. This decomposition does not have a causal interpretation since the exogeneity assumption $\mathbb{E}(\upsilon \mid T, Z, X) = 0$ may be violated even if the base specification (5) is identified. Nevertheless, it gives an indication of the importance of potential mediators. The contribution of variable k can be quantified as $\hat{\delta}_k = \hat{\Gamma}_k \hat{\beta}_k$; where $\hat{\Gamma}_k$ represents the effect of the intervention on mediator k, and $\hat{\beta}_k$ is the estimate for this variable in specification (6).

As mediators Z, we consider all trajectories which may precede the outcome Y. Thus, when the outcome is top income, we consider trajectories going through primary school performance (Top GPA yes/no), secondary schooling (yes/no) and a high-ranking occupation (yes/no). There are thus 8 trajectories, and we use the combination N, N, N (no top GPA, no secondary schooling, no high-ranking occupation) as the reference category.[18]

Figure 4 visualises the results for top income, regression results are in Appendix Table **??**).[19] The X axis shows estimates of Γ_ks, i.e. by how much the intervention increased the probability of observing a certain mediator outcome. The Y axis shows estimates of the βs, capturing the association between the mediator and top income. Thus, the area resulting from an interaction of these two estimates will be the estimate of the contribution of that mediator to the overall treatment effect, $\hat{\delta}_k = \hat{\Gamma}_k \hat{\beta}_k$. As reported above, the effect of the intervention on the probability of a top quintile income is 0.08, denoted "Total effect". The results suggest that the trajectory YYY, signifying top GPA, secondary enrolment and high-ranking occupation, is responsible for half of the total effect on earnings, or 4% points. This is due to the intervention increasing the probability of entering this trajectory by 6.5% points, and this trajectory being associated with a probability of being in the top income quintile of 62% points. Second place is taken by the trajectory YNY, which signifies having a top GPA and a high-ranking occupation,

[18]By construction, mediator pathways are correlated: if you follow path a, you do not follow path b. But this does not bias the estimates: the paths are exhaustive and mutually exclusive.

[19]We use top income instead of log earnings here, because the estimates for β are bounded between 0 and 1 and therefore more convenient to present graphically. Appendix Table **??** shows results for log earnings.

but no secondary schooling. This trajectory is associated with an increase of 42% points in the probability of earning a top income, and the intervention increases the probability of entering this trajectory by 2.2% points. The contribution by this trajectory is thus 0.9% points or 11% of the total effect. None of the other trajectories make quantitatively meaningful contributions to the total effect. Appendix Figure **??** presents the corresponding results for the intermediate outcomes secondary schooling and high-ranking occupation. They show that 69% of the overall effect on secondary schooling may be attributed to top GPA, and that the trajectory combining top GPA and secondary schooling explains 39% of the overall effect of treatment on high ranking occupation, followed by secondary schooling graduates who did not have a top GPA. Results for men are available in Appendix Tables **??** and **??**.

7 Robustness Checks

We perform a suite of robustness checks. We investigated alternative measures of intervention exposure, learning that exposure at 0-3 months is most effective in modifying outcomes. We show that tests of balance on baseline variables, tests for pre-trends, a placebo test and randomisation inference support our identification strategy. We show that attrition is only differential by treatment status in the school sample. To account for this, we re-estimate labour market outcomes on the school sample and, we also show that the results hold when we estimate Lee bounds. The earnings results are not sensitive to tests for selective survival. The significance of the estimates is robust to clustering at the health district level. Dropping the two cities, and dropping covariates, does not cause meaningful changes. The results are similar if we anchor school grades to income or education, as they are if we include uncertain matches in the school data. The model specifications and results are in Appendix **??**.

8 Concluding Discussion

Using unique longitudinal data in which individual outcomes are observed at different stages of the life course, we identify large impacts of a universal infant care intervention on school and labour market outcomes. A crude estimate of the internal rate of return suggests that the trial was highly cost-effective, and it was successfully scaled up following the short trial period that we analyse. Our findings are of contemporary relevance given that poor health and nutrition and deficient early childhood care are estimated to be causing about 200 million children under the age of 5 to fail to attain their cognitive potential, and that this has been identified as a key factor in the intergenerational transmission of poverty (Grantham-McGregor et al., 2007).

Intervention effects are highly correlated across outcomes, implying that it is largely the same individuals who drive the various effects. With the caveat that it is only descriptive, our mediator analysis suggests that cognitive attainment, secondary schooling and a high-ranking occupation contributed significantly to the increase in adult earnings. The analysis also highlights the importance of (a) looking at the distribution of test scores and earnings and (b) the relevance of institutional capacity and demand conditions. It shows that population health improvements can lead to a demand for higher education, and that the impact of infant health interventions on earnings will tend to be larger when the demand for skilled workers is rising. Our results also highlight that gender segmentation in the labour market, which may lead to differential demand growth for male and female labour, may contribute to understanding the tendency for infant health interventions to have different impacts on the earnings of men and women.

In our setting, as will be the case in many developing countries today, secondary school places were limited and entry was competitively determined. Although intervention effects on boy test scores were stronger on average, it mattered that treatment effects on the chances of scoring in the top quintile were significantly greater for girls. This led to intervention effects on girls progressing into secondary schooling with no corresponding effect for boys. Our analysis suggests that treated boys were thereafter on a lower trajectory than treated girls. Leveraging variation in the share of treated girls in secondary school catchment areas we demonstrate displacement of boys. In addition, we show that when these cohorts emerged onto the labour market, the movement of high-skilled women into high-earning sectors was facilitated by growth in labour demand for skilled women, a large part of which was in woman-friendly public sector jobs created by expansion of the welfare state. This said, our results do not suggest that early health investments in boys are unimportant. There are positive effects on school test scores and impacts on earnings may be realised in a setting where school capacity and market demand are more favourable. Even in our setting, the cognitive gains among boys may have had a positive influence in domains we do not measure.

Our main contribution, delineated above, is to illuminate the mechanisms linking infant health to adult earnings. In this process we argument various strands of the literature. First, we contribute to evidence demonstrating that early life health interventions can have a causal impact on cognitive attainment (Figlio et al., 2014; Black et al., 2007; Bharadwaj et al., 2013; Bhalotra and Venkataramani, 2013). Second, we contribute to a scarce literature providing evidence that cognitive performance and higher education contribute to earnings. Pre-school programmes such as Project STAR and the Perry intervention raised long term earnings by generating improvements not in cognitive skills but, instead, in health and non-cognitive skills (Chetty et al., 2011; Heckman et al., 2013; Baker et al., 2018). A vast body of research documents long run benefits of early life health interventions on earnings (Almond and Currie, 2011; Falk and Kosse, 2016). While it is implicit that the intervening mechanism is human capital accumulation, there is limited evidence on the importance of cognitive skills in this process. This is also the case for the closely related study by Bütikofer et al. (2019) who examine the long-run impacts of a nationwide mother-baby programme rolled out over decades in Norway leading to higher earnings on average.[20]

In a recent review Almond et al. (2017) argue that the effects of the early life environment on long run outcomes are often heterogeneous,*"reflecting differences in child endowments, budget constraints, and production technologies"*. We show that opportunities for skill accumulation and jobs matter. By arguing that these factors create a divergence in earnings returns to an early life intervention, we contribute to a literature that has highlighted gender differences in returns (Baird et al. (2016); Bhalotra and Venkataramani (2013); Garcia et al. (2018); Molina (2020); Bobonis et al. (2006); Maluccio et al. (2009); Maccini and Yang (2009); Field et al. (2009)).

[20]Our work also relates to the studies by Hjort et al. (2017) and Wüst et al. (2018) who show that exposure to a similar infant intervention led to reduced adult deaths from infections, cancer and cardiovascular disease.

References

Acemoglu, D., D. H. Autor, and D. Lyle (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of political Economy 112*(3), 497–551.

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

Aizer, A., J. Currie, P. Simon, and P. Vivier (2018). Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics 10*(1), 307–41.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–172.

Almond, D., J. Currie, and V. Duque (2017). Childhood circumstances and adult outcomes: Act ii. *Forthcoming Journal of Economic Literature*.

Almond, D., B. Mazumder, and R. Van Ewijk (2014). In utero ramadan exposure and children's academic performance. *The Economic Journal 125*(589), 1501–1533.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association 13*(6), 949–997.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics 118*(4), 1279–1333.

Baird, S., J. H. Hicks, M. Kremer, and E. Miguel (2016). Worms at work: Long-run impacts of a child health investment. *The Quarterly Journal of Economics 131*(4), 1637–1680.

Baker, M., J. Gruber, and K. Milligan (2018). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy Forthcoming*.

Bång, J. (2001). The Returns to Education - Using Data Retrieved from the Swedish National Census of 1930. *mimeo, Department of Economics, Uppsala University*.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. *Science 335*(6068), 582–586.

Benjamini, Y., A. Krieger, and D. Yekutieli (2006). Adaptive linear step-up false discovery rate controlling procedures. *Biometrika 93*(3), 491–507.

Bhalotra, S., M. Fernandez-Sierra, and A. S. Venkataramani (2018). Women's labour force participation and the gender wage gap across the distribution of tasks.

Bhalotra, S., M. Karlsson, and T. Nilsson (2017). Infant health and longevity: Evidence from a historical intervention in sweden. *Journal of the European Economic Association*, jvx028.

Bhalotra, S. R. and A. S. Venkataramani (2013). Cognitive development and infectious disease: Gender differences in investments and outcomes. *IZA Discussion Paper* (7833).

Bharadwaj, P., K. V. Løken, and C. Neilson (2013). Early life health interventions and academic achievement. *American Economic Review 103*(5), 1862–1891.

Bitler, M., T. Domina, and H. Hoynes (2016). Experimental Evidence on Distributional Effects of Head Start. Mimeo, UC Davis.

Björklund, A. and C. Kjellström (1994). Avkastningen på utbildning i sverige 1968 till 1991.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 93–120.

Bobonis, G. J., E. Miguel, and C. Puri-Sharma (2006). Anemia and school participation. *Journal of Human resources 41*(4), 692–721.

Borgen, N. T. (2016). Fixed effects in unconditional quantile regression. *The Stata Journal 16*(2), 403–415.

Borusyak, K., P. Hull, and X. Jaravel (2018). Quasi-experimental shift-share research designs. Technical report, National Bureau of Economic Research.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque (2009). Adult height and childhood disease. *Demography 46*(4), 647–669.

Bütikofer, A., K. V. Løken, and K. G. Salvanes (2019). Infant health care and long-term outcomes. *Review of Economics and Statistics 101*(2), 341–354.

Cattan, S., G. Conti, and C. Farquharson (2019). Workforce quality in early years intervention: Evidence from a large-scale home-visiting program.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Clarke, D., G. Cortés Méndez, and D. Vergara Sepúlveda (2018). Growing together: Assessing equity and efficiency in an early-life health program in chile.

Coles, M. G. and M. Francesconi (2017). Equilibrium search and the impact of equal opportunities for women. *Journal of Political Economy*.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2018). The" end of men" and rise of women in the high-skilled labor market. Technical report, National Bureau of Economic Research.

De Chaisemartin, C. (2012). Fuzzy differences in differences. Pse working papers, HAL.

de Chaisemartin, C. and X. d'Haultfoeuille (2019). Two-way fixed effects estimators with heterogeneous treatment effects. Technical report, National Bureau of Economic Research.

Deuchert, E., M. Huber, and M. Schelker (2016). Direct and indirect effects based on difference-in-differences with an application to political preferences following the vietnam draft lottery.

Deverman, B. E. and P. H. Patterson (2009). Cytokines and Cns Development. *Neuron 64*, 61–78.

Dhamija, G. and S. Gitanjali (2019). Impact of early life interventions on later life outcomes: Evidence from india's integrated child development services.

Dippel, C., R. Gold, S. Heblich, and R. Pinto (2017). Instrumental variables and causal mechanisms: Unpacking the effect of trade on workers and voters. Technical report, National Bureau of Economic Research.

Elsby, M. W., B. Hobijn, and A. Sahin (2010). The labor market in the great recession. Technical report, National Bureau of Economic Research.

Engle, P. L., M. M. Black, J. R. Behrman, M. Cabral de Mello, P. J. Gertler, L. Kapiriri, R. Martorell, and M. E. Young (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet 369*(9557), 229–242.

Eppig, C., C. L. Fincher, and R. Thornhill (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society of London B: Biological Sciences 277*(1701), 3801–3808.

Falk, A. and F. Kosse (2016). Early childhood environment, breastfeeding and the formation of preferences.

Field, E., O. Robles, and M. Torero (2009). Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics 1*(4), 140–69.

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

Finch, C. E. and E. M. Crimmins (2004). Inflammatory exposure and historical changes in human life-spans. *Science 305*(5691), 1736–1739.

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

Fischer, M., M. Karlsson, T. Nilsson, and N. Schwarz (2019). The long-term effects of long terms. compulsory schooling reforms in sweden. *Forthcoming in Journal of European Economic Association*.

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

Fredriksson, V., S. Marklund, G. Sivgard, and M. Widen (1971). Svenska folkskolans historia, sjätte delen. skolutvecklingen 1942–1962.

Garcia, J. L., J. J. Heckman, and A. L. Ziff (2018). Gender differences in the benefits of an influential early childhood program. *European economic review 109*, 9–22.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Goldin, C. (1988). Marriage bars: Discrimination against married women workers, 1920's to 1950's.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018). Bartik instruments: What, when, why, and how. Technical report, National Bureau of Economic Research.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.

Gorna, R., N. Klingen, K. Senga, A. Soucat, and K. Takemi (2015). Women's, children's, and adolescents' health needs universal health coverage. *The Lancet 386*(10011), 2371–2372.

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

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J. and R. Pinto (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric reviews 34*(1-2), 6–31.

Hinnerich, B. T. and P. Pettersson-Lidbom (2014). Democracy, redistribution, and political participation: Evidence from sweden 1919–1938. *Econometrica 82*(3), 961–993.

Hjort, J., M. Sølvsten, and M. Wüst (2017). Universal investment in infants and long-run health: Evidence from denmark's 1937 home visiting program. *American Economic Journal: Applied Economics 9*(4), 78–104.

Huber, M., M. Lechner, and A. Strittmatter (2017). Direct and indirect effects of training vouchers for the unemployed. *Journal of the Royal Statistical Society: Series A*.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the chicago public schools. *Journal of public Economics 89*(5-6), 761–796.

Killewald, A. and J. Bearak (2014). Is the motherhood penalty larger for low-wage women? a comment on quantile regression. *American Sociological Review 79*(2), 350–357.

Leeuwen, v. M., I. Maas, and A. Miles (2002). Hisco: Historical international standard classification of occupations. *Leuven University Press*.

Levine, P. B. and D. Schanzenbach (2009). The impact of children's public health insurance expansions on educational outcomes. In *Forum for Health Economics & Policy*, Volume 12. De Gruyter.

Lindgren, K.-O., S. Oskarsson, and C. T. Dawes (2014). Can political inequalities be educated away? Evidence from a Swedish school reform. Technical report, Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Lindgren, K.-O., S. Oskarsson, and M. Persson (2019). Access to education and political candidacy: Lessons from school openings in sweden. *Economics of Education Review 69*, 138–148.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–26.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein (2009). The impact of improving nutrition during early childhood on education among guatemalan adults. *The Economic Journal 119*(537), 734–763.

Molina, T. (2020, 03). Pollution, Ability, and Gender-Specific Investment Responses to Shocks. *Journal of the European Economic Association*. jvaa005.

Nowakowski, R. S. (2006). Stable neuron numbers from cradle to grave. *Proceedings of the National Academy of Sciences of the United States of America 103*(33), 12219–12220.

Pettersson-Lidbom, P. (2015). Midwives and Maternal Mortality: Evidence from a Midwifery Policy Experiment in Sweden in the 19th Century.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (2012). Human capital investment and the gender division of labor in a brawn-based economy. *The American Economic Review 102*(7), 3531–3560.

Saaritsa, S. and A. Kaihovaara (2016). Good for girls or bad for boys? schooling, social inequality and intrahousehold allocation in early twentieth century finland. *Cliometrica 10*(1), 55–98.

SCB (1972). Population and housing census 1970, part 1. population in communes and parishes, etc. *National Bureau of Statistics*.

Sievertsen, H. H. and M. Wüst (2017). Discharge on the day of birth, parental response and health and schooling outcomes. *Journal of health economics 55*, 121–138.

SOU (1935). Kungl. medicinalstyrelsens utlåtande och förslag angående förebyggande mödra och barnavård. *Statens offentliga utredningar 1935:19*.

Stanfors, M. (2003). *Education, labor force participation and changing fertility patterns*, Volume 22. Lund studies in economic history.

Statistics Sweden (1947). Kommunala valen. år 1946. Stockholm.

Stenhoff, G. (1934). Försöksverksamhet beträffande för- och eftervård vid barnsbörd. *Tidskrift för Barnavård och Ungdomsskydd 3*, 99–101.

Sundén, A. (2006). The swedish experience with pension reform. *Oxford Review of Economic Policy 22*(1), 133–148.

UNESCO (2014). *Teaching and learning: Achieving quality for all*, Volume 2013/4. UNESCO Global Monitoring Report.

Wüst, M., E. L. Mortensen, M. Osler, and T. I. Sørensen (2018). Universal infant health interventions and young adult outcomes. *Health economics 27*(8), 1319–1324.

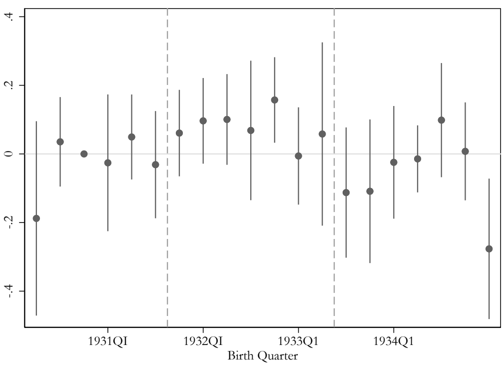
Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal 125*(588), F347–F371.

TABLE 1. RESULTS FOR OUTCOME INDICES: MULTIPLE HYPOTHESIS TESTING ADJUSTMENT

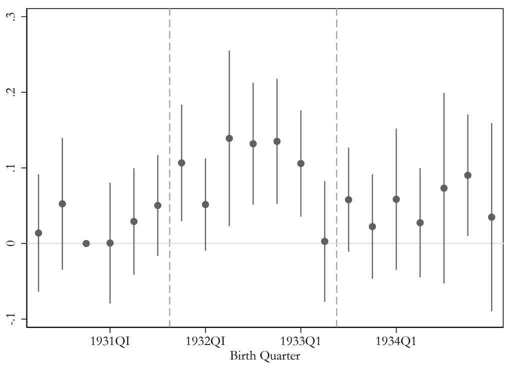
	Females		Males	
	N	Estimates	N	Estimates
A. Age 7 Index				
	5,830	0.0286	6,059	-0.0437
SE		0.120		0.084
p val		0.811		0.604
BH p val		1.000		1.000
B. Age 10 Index				
	10,298	0.1319	10,617	-0.0347
SE		0.051		0.048
p val		0.011		0.473
BH p val		0.029		1.000
C. Adult Index				
	10,301	0.0764	10,619	-0.0072
SE		0.022		0.018
p val		0.001		0.694
BH p val		0.005		1.000

Each coefficient is estimated using specification (1) with local trends. Standard errors are clustered at the parish level. 'p val' present conventional p-values; 'BH p val' presents p-values controlling the false discovery rate following Benjamini et al. (2006).

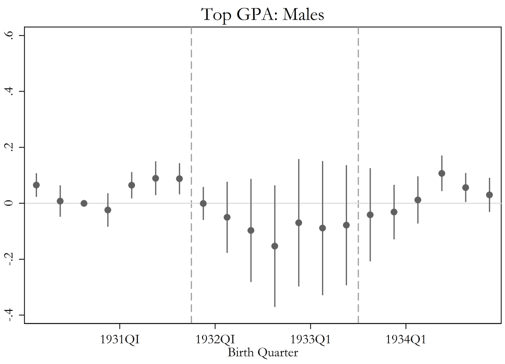
FIGURE 1. Event Studies



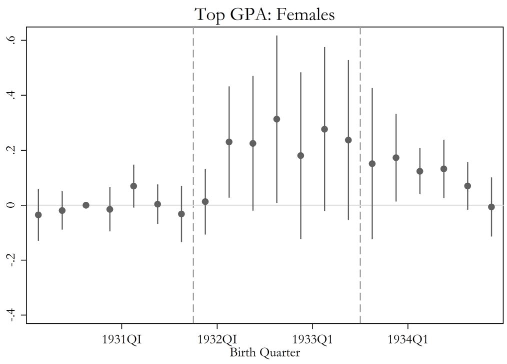
(a) Age 10 Index-Girls



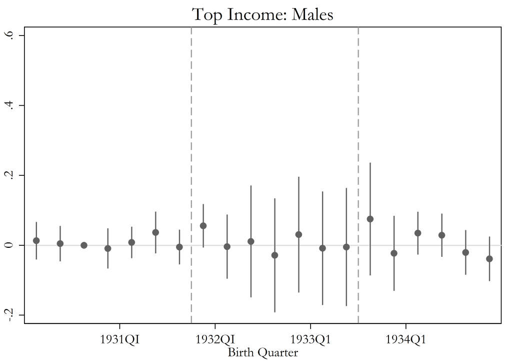
(b) Adult Index-Women



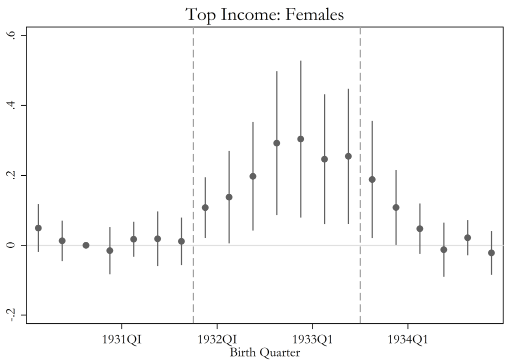
(c) Top GPA-Boys



(d) Top GPA-Girls



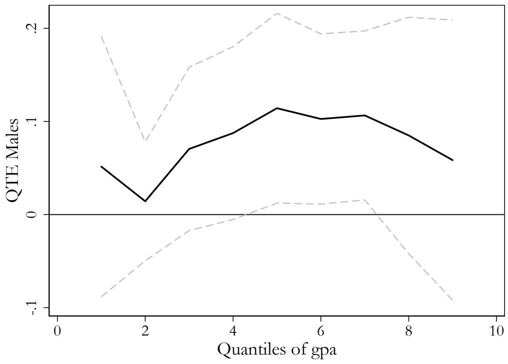
(e) Adult Income-Men



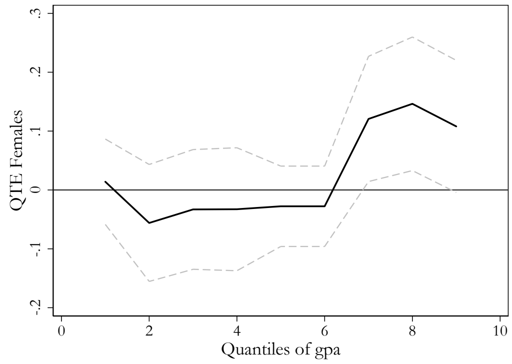
(f) Adult Income-Women

The upper panel (figures a and b) presents event studies for outcome indices for females. The indices are defined in section 4.1. The middle panel (figures c and d) presents event studies for Top GPA in grade 4 for boys and girls. The lower panel (figures e and f) presents event studies for Top Income for women and men. The vertical dashed lines signify the eligibility period of the infant care trial. 90% confidence intervals.

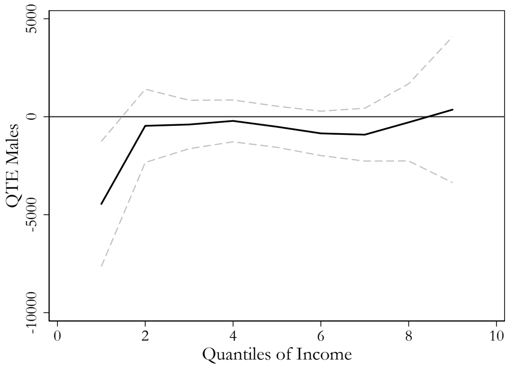
FIGURE 2. QUANTILE REGRESSION: GPA IN GRADE 4 AND INCOME BY GENDER



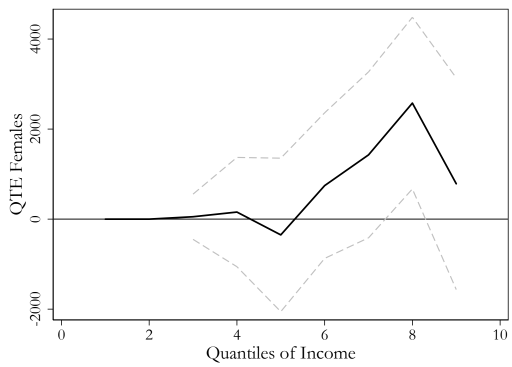
(a) GPA-Boys



(b) GPA-Girls



(c) Income-Men



(d) Income-Women

The upper panel (figures a and b) presents quantile regressions for GPA in grade 4. The lower panel (figures c and d) presents quantile regressions for income in adulthood. Covariates included are a dummy indicating twin births, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, Parish FE and QOB$\times$YOB FE. 90% confidence intervals included.

TABLE 2. COGNITIVE PERFORMANCE IN PRIMARY SCHOOL, GRADE 4

	Girls				Boys			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)
Top GPA	6,561	0.227	0.1000*	0.1243*	6,707	0.116	0.0400	0.0275
			(0.059)	(0.071)			(0.033)	(0.028)
GPA	6,561	0.098	0.0410	0.0617	6,707	-0.200	0.1213**	0.1084
			(0.049)	(0.054)			(0.057)	(0.072)
Math	6,554	0.025	-0.0535	-0.0217	6,688	-0.082	0.0193	0.0317
			(0.051)	(0.056)			(0.079)	(0.091)
Reading	6,536	0.120	0.0832	0.0902	6,687	-0.241	0.1823***	0.1649**
			(0.057)	(0.066)			(0.064)	(0.082)
Writing	6,536	0.150	0.0859	0.1068	6,692	-0.275	0.1645**	0.1291*
			(0.081)	(0.094)			(0.064)	(0.072)
Religion	6,549	0.088	0.0160	0.0654	6,689	-0.184	-0.0222	0.0247
			(0.052)	(0.066)			(0.097)	(0.096)
Parish FE			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓
School FE			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓
Length of Schoolyear			✓	✓			✓	✓
Schoolform			✓	✓			✓	✓
Parish Trends				✓				✓

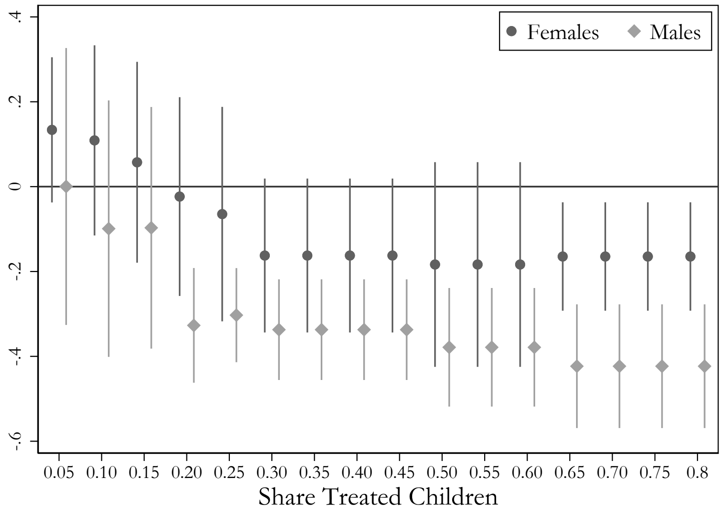
*** p <0,01; ** p <0,05; * p <0,1. Standard errors are clustered at the parish level. Covariates included are twin births, old (>35 years) and young (<20) mothers, married women, hospital birth, and exposure to the maternal intervention. 'Mean' refers to the mean value of the outcome variable before the intervention. 'QOB×YOB effects' are quarter-of-birth dummies. 'Parish FE' are fixed effects for the birth parish. 'SES effects' are fixed effects for the occupational group of the household head. 'Length of schoolyear' are fixed effects controlling for reforms on the length of the school year. 'Schoolform' are fixed effects controlling for the school form, see Section 2.2.

TABLE 3. SECONDARY SCHOOLING, EARNINGS AND EMPLOYMENT

	Women				Men			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)
Primary	10,105	0.675	-0.0087	-0.0011	10,369	0.725	0.0280	0.0367
			(0.032)	(0.026)			(0.023)	(0.025)
Dropout	10,105	0.126	-0.0196	-0.0277	10,369	0.101	0.0131	-0.0167
			(0.023)	(0.023)			(0.029)	(0.027)
Secondary	10,105	0.198	0.0353**	0.0350**	10,369	0.172	-0.0468	-0.0289
			(0.016)	(0.014)			(0.029)	(0.021)
Top Income 1970	10,307	0.244	0.0655***	0.0788***	10,613	0.210	-0.0445	-0.0361
			(0.022)	(0.028)			(0.034)	(0.028)
Log Income	10,307	8.990	0.1204*	0.1947***	10,613	10.222	-0.0596	-0.0464
			(0.063)	(0.066)			(0.037)	(0.036)
Log Pensions (age 71)	8,284	11.609	0.0293	0.0711***	7,680	11.995	-0.0400**	-0.0400*
			(0.019)	(0.015)			(0.017)	(0.020)
Working Parttime	10,256	0.265	-0.0325	-0.0244	10,466	0.019	-0.0077	-0.0049
			(0.030)	(0.033)			(0.007)	(0.007)
Working Fulltime	10,256	0.370	0.0607*	0.0760**	10,466	0.925	-0.0052	-0.0061
			(0.031)	(0.037)			(0.014)	(0.015)
Municipal	10,256	0.238	0.0377*	0.0488**	10,466	0.092	0.0012	0.0102
			(0.020)	(0.020)			(0.014)	(0.016)
Federal	10,256	0.051	0.0306***	0.0339**	10,466	0.111	-0.0053	-0.0077
			(0.012)	(0.014)			(0.019)	(0.019)
Parish FE			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓
School Reforms			✓	✓			✓	✓
Parish Trends				✓				✓

*** p <0,01; ** p <0,05; * p <0,1. See notes of Table 2. ‘School reforms’ refers to the extension of compulsory schooling and length of school year reforms.

FIGURE 3. SHARE OF TREATED CHILDREN: SECONDARY SCHOOLING



Each dot refers to treatment effects on secondary schooling completion when the share of treated children in the catchment area is at least the share of treated children noted on the x-axis. 90% confidence intervals.

TABLE 4. TREATMENT EFFECT HETEROGENEITY BY BARTIK INSTRUMENT FOR SKILLED WORKERS, ADULT INDEX

	Females (N=10,301)		Males (N=10,619)	
	(1)	(2)	(3)	(4)
Treated × Duration Eligibility	0.0721***	0.0747***	-0.0147	-0.0142
	(0.022)	(0.021)	(0.016)	(0.016)
Treated × Own Skilled Bartik	0.0070	0.0030	0.0366**	0.0366**
	(0.051)	(0.051)	(0.018)	(0.018)
Own Skilled Bartik	0.0409	0.0453	-0.0003	0.0002
	(0.039)	(0.038)	(0.009)	(0.009)
Duration Eligibility × Own Skilled Bartik	-0.0309**	-0.0311**	-0.0216*	-0.0224*
	(0.013)	(0.013)	(0.012)	(0.012)
Treated × Duration Eligibility × Own Skilled Bartik	0.0581***	0.0587***	0.0169	0.0196
	(0.018)	(0.018)	(0.017)	(0.019)
Treated × Other Skilled Bartik		0.0229		0.0352
		(0.017)		(0.040)
Other Skilled Bartik		-0.0321***		-0.0178
		(0.010)		(0.021)
Duration Eligibility × Other Skilled Bartik		0.0161		0.0014
		(0.014)		(0.009)
Treated × Duration Eligibility × Other Skilled Bartik		-0.0105		-0.0015
		(0.019)		(0.012)
Parish FE	✓	✓	✓	✓
QOB×YOB FE	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓
School Reforms	✓	✓	✓	✓
Parish Trends	✓	✓	✓	✓

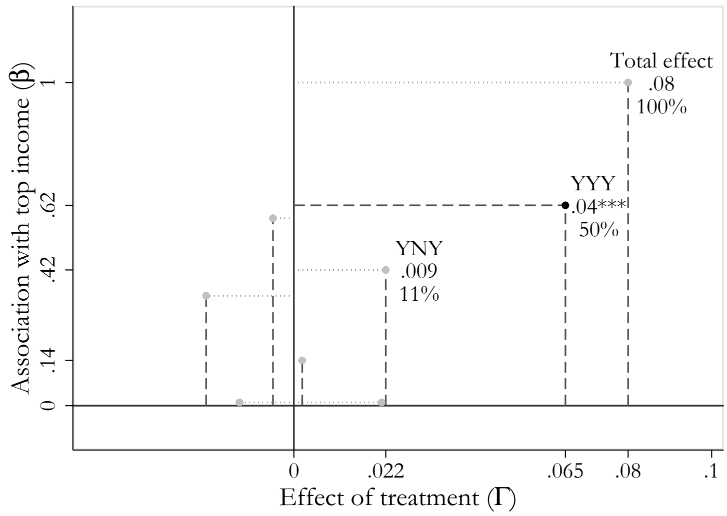
*** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. The Bartik index is defined at the cohort-parish level for skilled workers of each gender. 'Own' is own-gender 'Other' is other gender, see Appendix **??** for definitions. See notes of Table 3 for a list of controls.

TABLE 5. Correlated Treatment Effects: Women

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome 1	Outcome 2	τ_1	τ_2	τ_Y^{uc}	τ_Y		$\text{corr}\,(\tau_{1i}, \tau_{2i})$	
Top GPA								
	Secondary	0.1055*	0.0519*	0.0337	0.0664***		0.7738	
		(0.062)	(0.027)		(0.024)	[0.5332	–	0.8426]
	High Occ	0.1044*	0.0631	0.0485	0.0856**		0.9848	
		(0.063)	(0.056)		(0.039)	[0.9524	–	0.996]
	Top Income	0.1044*	0.0837*	0.0465	0.0704*		0.5420	
		(0.063)	(0.050)		(0.041)	[0.1697	–	0.6484]
Secondary								
	High Occ	0.0396**	0.0815**	0.0276	0.0458***		0.6121	
		(0.017)	(0.038)		(0.014)	[0.2426	–	0.7177]
	Top Income	0.0396**	0.0649**	0.0212	0.0392***		0.5825	
		(0.017)	(0.033)		(0.013)	[0.2213	–	0.6857]
High Occ								
	Top Income	0.0817**	0.0650**	0.0376	0.0568**		0.6005	
		(0.038)	(0.033)		(0.024)	[0.2748	–	0.6936]

τ_Y^{uc}: benchmark value, uncorrelated effects, see Appendix **??** for a derivation; τ_1: treatment effect outcome 1; τ_2: treatment effect outcome 2; τ_Y: joint treatment effect for interacted outcome 1×2; $\text{corr}\,(\tau_{1i}, \tau_{2i})$: correlation coefficient between treatment effects. Bounds for alternative assumptions in square brackets.

FIGURE 4. GELBACH MEDIATION, WOMEN: TOP INCOME



The figure shows how different trajectories defined by a) top GPA (Y/N), b) secondary schooling enrollment (Y/N) and c) high-ranking occupation (Y/N) contribute to the estimate of the overall effect of the intervention on top income. The X axis measures the effect of the intervention on the mediator/outcome and the Y axis measures the association between the mediator and the main outcome (top income). Dashed lines represent significance at the 1% level, dash-dotted lines represent significance at the 10% level, and dotted lines represent insignificant estimates. See Appendix Table **??** for regression results.