Long-Run and Multi-Generational Impacts of Investments in Child Health: Evidence from a Government Trial

Jennifer L. Baker* Lise G. Bjerregaard* Christian M. Dahl†

Torben S. D. Johansen† Emil N. Sørensen‡ Miriam Wüst§

Work in Progress — Version: July 10, 2022.

Abstract

We transcribe handwritten nurse records on Danish children born in the 1960s and link them to administrative population data to study how preventive health care for toddlers affects them in the long run. We exploit variation from a government trial at scale, extending a universal home visiting program to not only cover infancy but two additional years for ten percent of Copenhagen children, those born on the first three days of the month. We find health improvements over the life course for treated children on measures of childhood health, an adult good health index, and short- and long-run mortality. While we do not find impacts on a combined education and labor market index, treated girls experience significant improvements in labor market outcomes likely related to improved health. For children with initial health disadvantages long-run health impacts are largest and may even extend to their own offspring. Our findings demonstrate that universal early-life health policies during childhood can contribute to the alleviation of adult health inequalities at very modest per-child costs.

*Center for Clinical Research and Prevention, Copenhagen University Hospital – Bispebjerg and Frederiksberg, Copenhagen, Denmark.

†Department of Economics, University of Southern Denmark.

‡The School of Economics, University of Bristol.

§Corresponding author: University of Copenhagen, Center for Economic Behavior and Inequality (CEBI), Department of Economics, Øster Farimagsgade 5, 1353 Copenhagen, email: miriam.w@econ.ku.dk.

We gratefully acknowledge financial support from the Independent Research Fund Denmark, grant 8106-00003B. Wüst acknowledges support from the Center for Economic Behavior and Inequality (CEBI) at the University of Copenhagen, financed by grant DNRF134 from the Danish National Research Foundation.

We thank the Copenhagen City Archive and especially Nicolai Eberholst for their support. In this project, we transcribe and use data from the "Copenhagen Infant Health Nurse Records (CIHNR)". The CIHNR was established in a collaboration between the Center for Clinical Research and Prevention, The Capital Region of Denmark, the University of Copenhagen and the University of Southern Denmark. The authors thank Thorkild I.A. Sørensen for his contributions that led to the establishment of this data source.

We thank Fane Groes, Peter Sandholdt, and seminar participants at CEBI (UCPH), the University of Zurich, Uppsala University, Royal Holloway, Universidad Autonoma de Barcelona, the NORC Demography Workshop (University of Chicago), Copenhagen Business School, Erasmus University Rotterdam, the 2022 Essen Health Conference, the 2022 ESPE conference and the Copenhagen Education Network for helpful comments.

1 Introduction

An established literature across disciplines documents the importance of early-life circumstances for the health and economic well-being of individuals over the life course (Forsdahl, 1979; Almond et al., 2018). While initially the negative impact of early-life insults to health and development has been center stage in this work, more recently, the role of health policies, their timing, targeting, and content has gained increased interest in economic research.[1] Motivated by theoretical models on early-life skill formation (Heckman, 2006; Attanasio, 2015), this work attempts to identify the role and relative importance of different inputs in child health and human capital production.

Existing empirical research on the long-run importance of early-life health policies falls broadly in two streams: First, randomized trials with high-intensity model programs such as the Nurse Family Partnership or the U.S. Perry Preschool Program have made a strong case for the considerable long-run returns to targeted investments in the health and development of disadvantaged children (Olds et al., 1986, 1998, 2019; Belfield et al., 2006; Heckman et al., 2010). Second, an observational literature often using large-scale administrative data has documented the long-run effects of initial access to early-life health policies on adult health and socioeconomic outcomes. Examples for this line of work include studies on the long-run impacts of access to early-life nutritional and income support (Hoynes et al., 2016; Bailey et al., 2020a; Barr et al., 2022), access to health insurance and thus care (East et al., 2022; Wherry et al., 2018; Goodman-Bacon, 2018; Miller and Wherry, 2019; Noghanibehambari, 2022), access to early education programs (Rossin-Slater and Wüst, 2020; Bailey et al., 2020b), and access to infant nurse home visiting and center care (Hjort et al., 2017; Bhalotra et al., 2017; Bütikofer et al., 2019; Hoehn-Velasco, 2021). Importantly, given large data requirements, these studies typically focus on aggregate-level variation in the introduction of or eligibility for early-life programs and thus intention to treat impacts.

[1] For studies on nutritional, disease, or income shocks in utero or early in life see, for example, (Almond, 2006; Almond et al., 2009; Schulz, 2010; Van den Berg et al., 2010)

This paper combines some of the strengths of these two lines of empirical work to study the long-run and multi-generational impacts of the design of early-life health policies: We exploit variation from a government randomized trial and use population administrative data on a range of long-run as well as second generation outcomes. Our study is unique, because we can link individual-level data on background characteristics, treatment exposure and outcomes. Moreover, rather than focusing on a targeted intervention for at-risk families, we exploit variation in a universal program for a general population of families.

We study a 1960s government trial in the Danish capital Copenhagen, which extended the duration of an existing home visiting program from providing preventive care to all infants the first year of life to also covering the second and third year for ten percent of all children, those born on the first three days of each month. In the first year program, nurses monitored infant health and development, counseled parents on investment decisions such as infant nutrition and preventive care/vaccination uptake, and, importantly, referred ill infants to other health care providers. In the follow-up program, assigned to Copenhagen resident families during the final first-year visits, nurses continued to focus on health: They monitored child health, encouraged parental health investments such as dental care, vaccines and healthy nutrition, advised on prevention of accidents, and referred ill children to follow-up treatments. Additionally, they advised parents on less health-centred topics such as the socio-emotional development of young children, parenting styles, and the take-up of childcare. On top of the on average between 12 and 14 first year nurse visits, families in the treatment group were offered three and two visits in the second and third year of the child's life, respectively.[2]

To make our study feasible, we create and combine several data sources. We use Danish administrative data for the 1977-2020 period providing a range of outcome measures for focal children, their mothers and offspring. A 1959-1961 cohort study allows us to measure intermediate childhood outcomes. Finally, to measure the assignment and implementation

[2]It is important to note, however, that in this period nurses played an instrumental role in the Danish health care system for referring children with identified health needs to additional care. Thus the suggested five visits by nurses very likely increased the number of overall health care provider contacts in the treated group.

of the trial and measure family background at the individual level (in the period predating administrative data at Statistics Denmark), we transcribe handwritten nurse records for the trial cohorts from the Copenhagen City Archives. To retrieve data from these historical documents with a complex table structure, we develop and use techniques for layout detection (Clinchant et al., 2018; Shen et al., 2021) and scene, optical character, and handwritten text recognition (Goodfellow et al., 2013; Bluche et al., 2014; Lee and Osindero, 2016; Bluche et al., 2017). Our approach allows us to transcribe millions of fields from the nurse records with high accuracy. At the same time, it is very general, making extension to other types of documents with a table structure, such as censuses, feasible.[3]

Our first stage results document that 57 percent of children born on the first three days of the month were assigned to the extended nurse program at age one. We measure treatment assignment by identifying the table structure (and the completeness of registrations) in the nurse records and our first stage holds across birth years and months, and across family and child characteristics likely observed by nurses. This finding across predetermined characteristics underscores the universal assignment of extended nurse care.[4]

In the long-run, we find that treated individuals score 4.4-5.0 percent of a standard deviation higher on a good health index, summarizing a set of hospital diagnoses and admission outcomes for individuals in their 30s through 50s. This impact is stronger for girls. Examining underlying health measures separately in this relatively young and thus healthy group of individuals, we find impacts that are positive across the board but only individually significant for measures of asthma and cancer diagnoses. At the same time, we find no average impacts of extended nurse visiting on a socioeconomic (SES) index summarizing education and labor market outcomes. This result conceals heterogeneity across gender and dimensions of our SES index: We find positive effects for girls and those are concentrated in labor market

[3]We use a unique Danish personal identifier to link all data sources.

[4]The share of non-compliance squares well with aggregate sources that show that initial compliance with the first-year nurse program was high but that around 30 percent of families discontinued their participation in the nurse program during the first year (due to mobility and some mortality). Thus exit out of the first-year program resulted in children not being eligible for treatment assignment (Copenhagen Statistical Office, various years).

dimensions likely susceptible to adult health (the share of time in employment during ages 30-50 and the probability receiving a disability pension at age 50). This finding is in line with impacts on the health index and may indicate impacts on labor market outcomes in the very long run (when individuals make decisions about labor market exit).

Another important dimension of health–not included in the good health index–is mortality. We find that treated children experience lower mortality risks in the short run (measured as ever being observed in our outcome data). Also in the long-run, we find that treated individuals are less likely to die (a seven percent decrease in the probability of death before or in the year 2020 at the relevant sample mean of 8.3 percent (ITT) or a 13 percent effect on treated individuals). This large effect points to potential larger health returns in the very long run. Additionally, given that we do not factor in mortality and thus exit from our outcome data in our main analyses for the health index, those are likely lower bounds (likely overstating the adult health in the control group).

Exploring mechanisms for our long-run findings, we rule out that maternal fertility or labor market decisions drive our results. We document some positive impacts on childhood health, in particular on early childhood height (among boys and fading out over childhood years) and on measures of infectious disease exposure and consumption of antibiotics, especially for girls. These results support our long-run results for asthma, which has been related to early childhood conditions, including nutritional inputs, exposure to passive smoking, exposure to inflammation and use of antibiotics (Gern et al., 1999; Burbank et al., 2017; Patrick et al., 2020). While our childhood health results suggest that nurses were successful in preventing illnesses and promoting some health investments in the family home, we do not detect impacts along other margins of health investments and other dimensions of child development. This last result may be due to nurses' qualifications, or the dosage and content of the developmental advice given by nurses in the second and third year program.

While policy decisions after the trial were not based on formal analyses of trial data, its existence allows us to also speak to a central question in contemporary policy debates: Should

well-child care beyond the first year of life be universally provided to all families or should we target efforts? In the latter case, can providers–such as nurses–easily screen families that likely benefit the most of prolonged care? Exploiting our data on pre-treatment family characteristics that are easily observed by nurses, we study dimensions of effect heterogeneity, among them initial health disadvantage.[5] Low birth weight children are much more positively impacted by extended nurse care than non-low birth weight children, scoring 24.3 percent of a standard deviation higher on the good health index and experiencing larger long-run survival benefits. Adding to this first order long-run impacts of extended nurse care for focal children, we finally study outcomes for the offspring of focal children. Given large individual and societal costs associated with poor health as early as at birth and its documented persistence across generations (Currie and Moretti, 2007; Kreiner and Sievertsen, 2020), intergenerational effects may be instrumental for an assessments of the cost effectiveness of the policy (and alternatively, targeted policies). We find some positive impacts on health at birth in the second generation, but only for the offspring of focal children with initial poor health.[6] These results are more suggestive than our first generation results (given small cell issues when analyzing rare outcomes such as low birth weight across generations in our sample).

Our results contribute to an extensive literature on the short- and long-run impact of early-life circumstances (Forsdahl, 1979; Barker, 1990; Currie and Almond, 2011; Almond et al., 2018). Importantly, we add to the still limited literature on the importance of positive childhood health interventions beyond the first year of life and on the importance of universal care (Wüst, 2022). Comparing our estimates to the literature on more targeted policies and on policies focusing on the in utero and first year period, we document smaller but relevant average health impacts of toddler preventive care. This care (an understudied margin so far) improved long-run health for treated children at very modest per-child costs (that we

[5]Importantly, heterogeneous impacts are, as we show, not driven by a differential first stage but may partly be driven by more intensive counseling after randomization into the treatment. In our data, we do not find evidence for more intensive treatments (e.g. more visits in the second and third year conditional on being in the treatment group) to disadvantaged children.

[6]As we discuss, we find limited evidence for (small) fertility responses predominantly at the intensive margin.

estimate at around 342 USD (2020 USD) for the combined program, i.e. the first year program covering all children and follow-up for a 10 percent subset). Likely mechanisms include improved parental health investments, the uptake of additional (preventive) care and timely treatments, and, as a result, improved childhood health. A final contribution is our movement beyond the margin of the long-run impacts of formal access to care, considered in many settings, among them "roll-out studies" in the Nordic countries (Wüst, 2012; Hjort et al., 2017; Bhalotra et al., 2017; Bütikofer et al., 2019; Rossin-Slater and Wüst, 2020). This work has used aggregated data on treatment exposure (because individual-level data on treatment exposure did not exist). In contrast, we exploit data on both individual background characteristics and treatment status made accessible through our application and advancement of automated transcription methods. These data allow us to analyze heterogeneity of impacts across easily-observed dimensions, suggesting that childhood health policies can help alleviate initial inequality, even in the case of a universal, low-intensity and low-cost preventive care program as the one studied here.

2 Institutional Background and the Copenhagen Trial

In Denmark, universal preventive care in the nurse home visiting program for families with infants dates back to 1937. The Danish National Board of Health (DNBH) centrally designed the program to combat high infant mortality rates of around six percent in the early 1930s. Exploiting the staggered introduction across Danish municipalities in the 1937-1949 period, earlier research has documented both short- and long-run health benefits of the introduction of the program (Wüst, 2012; Hjort et al., 2017).[7] Guidelines for the original program mandated at least ten visits in the first year of the child's life where nurses advised new parents on topics such as infant nutrition and infant care, monitored infant and family health, and

[7]The introduction of the program decreased infant mortality, with the program saving 5-8 lives per 1000 live births. In the long run, program exposure of the 1930s and 1940s cohorts impacted their survival and cardiovascular health: Exposed individuals are less likely to die and less likely to have a cardiovascular disease diagnosis in the 45-64 years age range (1.3-2.8 percent decrease in the probability of being diagnosed at the control mean of 26.6 percent).

referred ill infants to general practitioners for treatment.

Throughout the 1950s and 1960s, in the light of societal changes and improvements of general living conditions, the DNBH planned an update of program guidelines. First, infant mortality had declined significantly since the introduction of the program (to around two percent in 1960 (Det Statistiske Departement, 1964)). A large share of this decline was in mortality from preventable causes (mainly infectious disease). In a new "low mortality environment", the DNBH emphasized the need to update the type of preventive health care provided to focus on broader health monitoring, the prevention of accidents and, importantly, the encouragement of relevant parental health investments such as healthy nutrition and habits, dental care and vaccinations. While these health services were already in place (vaccinations) or expanding (free child dental care), broad uptake was still a central concern (Copenhagen Archives, various years). Second, concurrent to changes in health needs, the ongoing expansion of formal daycare and the increase in female employment inspired the DNBH to also direct attention to new topics related to family life, such as parenting styles and the socio-emotional development of children. Third, as health and social inequality across different parts of the country remained or even increased, and given that nurse home visiting was still a voluntary municipal program, the DNBH planned to make the program mandatory after a general overhaul. Thus the DNBH initiated work to experiment with new design features in the program (Copenhagen Archives, various years).

As a central element in this strategy—discussed in a report published in 1970 (DNBH, 1970)—the DNBH initiated a set of experiments with extended nurse visits beyond the first year of the child's life, among them the Copenhagen trial.[8] This trial with extended nurse care started with children born on January 1, 1959. Nurses offered the additional follow-up to families during one of the final scheduled first year visits, i.e. inclusion in the trial was

[8] Aside from the Copenhagen trial, a couple of smaller initiatives were implemented in municipalities across the country, none of which were as large and carefully designed as the Copenhagen trial: Municipalities in the counties of Holbæk (from 1960), Aarhus (from 1962), and Esbjerg (from 1961), extended nurse care beyond the first year of life for all families or subsets of families with identified health and social risk factors. None of the other initiatives used a randomization strategy.

conditional on residence in Copenhagen and participation in the first year nurse program.[9] The treatment group in the trial included all children born on the first three days of the month as well as all twins. Importantly, the government trial in Copenhagen was conducted largely without additional national or municipal funding and thus without additional manpower. To compensate nurses, who had to perform additional visits to around 10 percent of the around 160 children in each nurse district, the number of first year visits was adjusted for all children: Nurses were instructed to reduce the number of first year visits for all families with one to two visit during the trial period from the pre-trial average of around 14-15 visits (Copenhagen Archives, various years).

Tables 1 and 2 illustrate, based on nurse records and instruction material preserved in the Copenhagen City Archives, the content of the first year and the extended nurse program in the Copenhagen trial: During the first year of life, visits and registrations by nurses focused particularly on infant health and development, infant nutrition, and family living conditions including economic conditions and the mother's labor force participation. During the final first year visit around child age 12 months, nurses recorded information related to vaccination coverage and health care usage.

[9]The head of the Copenhagen nurse program, Dr. Biering-Sørensen, and his scientific approach towards developing the nurse program played a central role in the organization of the Copenhagen trial, as documented in archive material (Copenhagen Archives, various years): First, during his time as program head, Copenhagen had already initiated a number of structured initiatives, including a WHO-sponsored examination of side effects of the calmette vaccine in infants, a project on pregnancy visits for mothers from one of the maternity wards in the city, and a project on documenting typical first-year health issues observed in a sample of children (those born on the first three days of each month). Thus the Copenhagen program was geared for implementing new elements into the existing program. Second, Biering-Sørensen focused on the quality of registrations in nurse records and on including structured data in nurse records that would be easy to use in research (as documented in earlier projects mentioned above). Thus the trial with extended nurse care was from the onset using suitable nurse records to be able to follow up on its results. Third, Biering-Sørensen's focus on scientific methods (choosing a random subset of families to be included in the trial and choosing to enroll all twins to potentially compare outcomes of mono-zygotic and di-zygotic twins) was visionary (even though power calculations were not part of the design of the program and the assignment mechanism). Additionally, archive material documents extensive preparation meetings for nurses prior to the trial and frequent discussions on practical aspects and the content of the trial during regular staff meetings. Finally, archive material illustrates that the trial was high on the political agenda: Copenhagen decision makers including Urban Hansen, deputy major for social affairs from 1956, explicitly supported and approved the trial.

Table 1 Content of the First Year Home Visiting Program in the Municipality of Copenhagen.

		Age of Child		
Topic	Examples of items in record	2 weeks	1, 2, 3, 4, 6, and 9 months	12 months
Family	Socioeconomic status, mother mental and physical health		✓	✓
Mother labor force	Employment status, childcare status		✓	✓
Nutrition				
a.	Infant feeding, number of meals	✓	✓	✓
b.	Duration of breastfeeding			✓
Child development				
a.	Smiles, lifts head, babbles, sits alone		✓	✓
b.	Height			✓
c.	Weight	✓	✓	✓
d.	Walks, #teeth, general health assessment, vaccination status, ever hospitalized			✓
Child sleeping conditions	Own bed		✓	✓

Notes: The table shows topics covered in the first year nurse program and example items, which nurses registered information on in the child's records. At each age, more than one nurse visit could be performed (depending on family needs) with an average of around 13 first year visits during the trial (Copenhagen Archives, nd). For each age-specific topic, nurse registrations were recorded at one of those visits.

Table 2 Content of the Home Visiting Program in the Municipality of Copenhagen, Second and Third Year Program Extension.

		Age of Child				
Topic	Examples for items in record	15 months	18 months	24 months	30 months	36 months
Health care take-up	GP consultation, dentist visits, vaccinations, diagnosed illnesses	✓	✓	✓	✓	✓
Nutrition	Vitamins, candy, teeth brushing	✓		✓		✓
Illnesses	Dyspepsia, lung illnesses, otitis, other, accidents	✓	✓	✓	✓	✓
Social development	Emotional problems, habit formation		✓		✓	✓
Parenting style, daycare	Strict, parent-child relation					✓
Child sleeping conditions	Own bed			✓		✓

Notes: The table shows topics covered in the second and third year nurse program and example items, which nurses registered information about in the child's records. The extended program offered five visits during the child's second and third year of life.

Visits during the second and third year of the child's life continued to have a strong focus on toddler health, as illustrated in Table 2. Nurses encouraged the take-up of (preventive) health care at general practitioners and dentists, and focused on the early identification of potential illnesses. In an environment with few scheduled interactions with other health professionals, nurses facilitated additional care and treatments. They informed about the importance of preventing accidents and about the benefits of health investments such as vaccinations and a healthy nutrition. Moreover, nurses were encouraged to monitor the socio-emotional development of children, as well as advise on parenting style and parent-child interactions.

3 Data Sources

We generate and combine three main data sources: First, to measure treatment and relevant family background characteristics, we use handwritten nurse records preserved in the Copenhagen City Archives for the 1959-1967 cohorts.[10] As the starting point, we use the full set of scanned paper nurse records. A subset of data from these records–the names of both parents and the child, the date of birth of parents and child, the sex of the child, as well as a set of nurse registrations for a sample of children–have been transcribed manually. Extending manual transcription beyond these data is not feasible due to scale. Instead, we apply a machine learning approach to collect the information registered in the nurse records in an automated fashion.

Second, the transcribed names and dates of birth data from the nurse records allow for an individual-level link of the records with administrative data on outcomes for the given cohorts, their parents, and their children at Statistics Denmark via the unique Danish personal identifier from the Central Person Register (established in 1968). Third, also using this personal identifier, we add supplementary data on childhood health and development from the Copenhagen Perinatal Cohort (CPC), a cohort study covering children bron in

[10]The time frame of the preserved records suggest that these were archived due to the trial.

1959-1961 in the largest hospital in Copenhagen (Rigshospitalet).

3.1 The Copenhagen Nurse Records: Transcription and Merge

Data Extraction from Handwritten Nurse Records Figure 1 depicts a typical scan of a multi-page nurse record for an infant in the treatment group. The nurse records, while hand-written, have a number of pages with standardized layout: Typically the first page contains a table with structured information on nurse observations during the first year of life. Additionally, in the top and bottom of this page, nurses registered details on family background (and first year health investments such as vaccine uptake). Registrations in the main table in the middle of this page are structured around the child's age at the specific visit were the registration took place (rows) and across different topics (columns) coded in standardized ways, such as "1/2" for "child can sit" or "1/2/3" for "type of child nutrition" (breastfeeding, partial breastfeeding, other). The second page in Figure 1 (the flip side of the record) contains cells for information on date, child age and weight and mode of nutrition at each performed visit (there were typically more visits than the ones that required registrations on the first page), and space for nurse observations in free text. Registrations on this page are less structured than registrations on the first page. Finally, the third page contains the "treatment table" with standardized registration options (rows) during second and third year visits (columns) and a flip side with free text. Thus the presence of this page and table in a child's record indicates that she was offered the extended program. Importantly, the treatment table is often but not always added as the third page of a record, i.e. the ordering of pages varies: The records vary in their total number of pages (1-20), with the vast majority of records having 2 or 4 pages.[11]

Analyzing data from the nurse records requires us to transform images (scanned records) into data, a high-dimensional prediction problem. We split our data extraction problem

[11]Instruction material preserved in the City Archive describes that nurses were to add additional pages to the record of relevant (treatment) children when offering the extended nurse program in one of the final first-year visits. Browsing the actual paper records, we have established that nurses typically stapled the page with the treatment table to the original record.

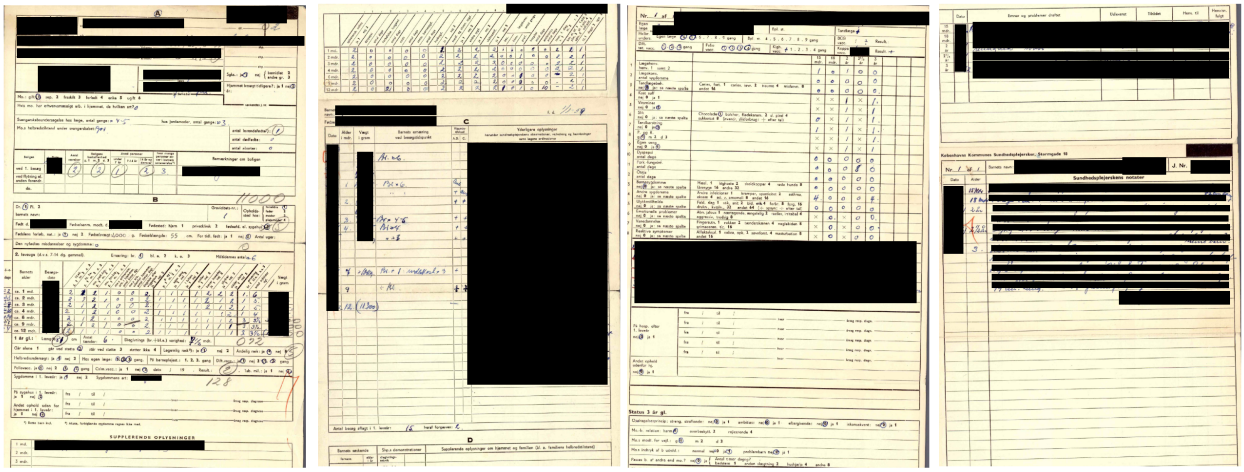


Fig. 1 Sample Nurse Record for a Copenhagen Child.

Notes: The pages depicted here come from a scanned nurse record. For confidentiality reasons, parts of the pages are blackened. The first page contains the table for first-year nurse registrations. The second page (flip side) is a page mainly for nurse comments in free text rather than predefined codes. The third page contains the table for second and third year registrations (the "treatment table"). The final page (flip side) allows for further nurse comments during the second and third year.

into three tasks that we describe in detail in Appendices D.1 and D.2. First, using an unsupervised clustering approach, we identify the presence of the treatment table (the table for standardized registrations in year two and three of the child's life). The resulting indicator of the presence of a treatment table in a given nurse record is our main measure of treatment assignment.[12] Second, using a neural network, we classify the content of the treatment table at the cell-level to create a compliance measure (a measure of whether the cell is filled out rather than its specific content). Third, also using neural networks, we transcribe the specific content of handwritten fields across different tables in the records. We apply the neural networks to cell-level images extracted from the records using table segmentation methods.

We transcribe a host of information registered by nurses alongside with the dates of the registration. The transcribed measures include information on the child (including birth weight and length, measures of infant weight and length, breastfeeding status, infant de-

[12]Our approach does not require training data, because we do not transcribe the specific content of the cells but identify the presence of the treatment table. However, we manually classify 4000 nurse records to assess the performance of our clustering method. We find that our assignment indicator predictor has precision and recall of unity.

velopment (smiles, lifts head, sits), infant nursery/daycare attendance status) and on the family (including its economic status, family "harmony" and cleanliness of the home, as well as measures for mothers' general physical and mental health status). The neural networks applied in these final steps need training data to transcribe cell content. We use the manually transcribed data for a subset of cells in the nurse records for this purpose. Moreover, to assess our approach, we also use a random subset of data for each manually transcribed field as test data to show that our transcription neural networks perform in the range of 96–100 percent for most table fields.[13]

Merge and Coverage of Nurse Records Merging the nurse records to other data sources proceeds in the steps illustrated in Figure 2: First, among the scanned original paper records, we have manually identified and excluded duplicate scans and some irrelevant documents (e.g., notice of a family moving) that have been archived together with nurse records.[14] The remaining 92,902 scanned documents constitute relevant nurse records.

Second, to link the nurse records to other data sources, we use the manually transcribed information on mothers', fathers' and children's names and dates of birth. These detailed data allow for a link of the records to the Danish Central Person Registry (CPR). The CPR contains a unique personal identifier (pnr) for each individual residing in Denmark in any year from 1968 on. Standard matching algorithms provide a CPR link for 88,808 (96 percent) of the relevant records. For the remaining 4,094 (four percent), the information on child and parent names and date of birth data does not result in a match with the CPR register.[15]

[13]We detail our assessment of the performance of the ML methods in Appendices D.1 and D.2. As our transcription approach relies on segmentation of cell-level images, this method fails for cases where the quality of a scan is particularly poor and the field as a consequence is intelligible. However, this loss depends on factors such as the quality of the scans and the condition of the source document, which are likely to be independent from treatment assignment. Thus the data that we lose here has no implications for the representativeness of our transcribed data, but only for precision in the sense of limiting our number of observations.

[14]Analyzing the dates on these documents, we find that those are not evenly distributed over the 1959-1967 period but predominantly come from the earliest years. This pattern may indicate that archiving was less stringent in the earliest years, i.e. that not only actual nurse records were kept and adding other types of files to the nurse archive.

[15]This failure of the match can be due to three reasons: The record information can be incomplete (names and birth dates can be missing on the record), making unique matches with the administrative data infeasible.

Third, after having excluded scans without a personal identifier, we link the nurse records to administrative data on relevant outcomes at Statistics Denmark. These data comprise all individuals born in the 1959-1967 period and observed in the administrative data at least once in the period 1977-2018. Moreover, the data include information on the parents, siblings, biological children, and first spouses of the focal individuals, if ever observed in the administrative data. We merge all but 808 nurse records with a valid identifier to the research data provided at Statistics Denmark. Given that these 808 individuals are observed in the 1968 CPR data as Danish residents, we document that the majority of non-linked records refer to individuals who left Denmark in the 1968-1977 period.[16]

In sum, a total of 4,902 relevant nurse records (5 percent) remain unmatched to our administrative outcome data at Statistics Denmark (and the match rate for the 1968 status is around 96 percent).[17] A final exclusion in our analyses relates to multiple births: Since all twins were offered the extended program, we exclude them from our analyses. Thus we exclude an additional 664 individuals (0.8 percent of the matched sample) with either the same mother and date of birth, or the same date of birth and no identified mother, from our final analysis sample (N=87,336).

Individuals may have died or emigrated prior to 1968, the year of establishment of the Danish personal identifier system. Finally, some of the hand-written documents and their scans are in too poor quality to identify names and dates of birth for the child and/or mother correctly and can therefore not be matched. We analyze the impact of non-matches (and thus potential selective emigration or mortality) in section 5.2.

[16]The CPR provides on request a status variable identifying the resident status of all individuals ever registered with a Danish residence since 1968. This status variable is usually not available fro the pre-1977 period in the Statistic Denmark research data. In our case, we cannot rule out that individuals with an emigration status died abroad. Thus the CPR status measure may understate deaths in the non-merged group.

[17]Appendix Figure A1 suggests that the share of unmatched nurse records (due to a missing identifier or missing outcome data at Statistics Denmark) is relatively stable across cohorts, with slightly higher match rates in younger cohorts. Across the treatment and control days of the month match rates are slightly higher for treatment days, as we document in our main analyses in Table 8. However, the difference is very small (a 0.4 percentage point higher merge probability at a baseline merge rate of 96 percent). We conclude that this difference likely reflects an impact of the extended nurse program on mortality.

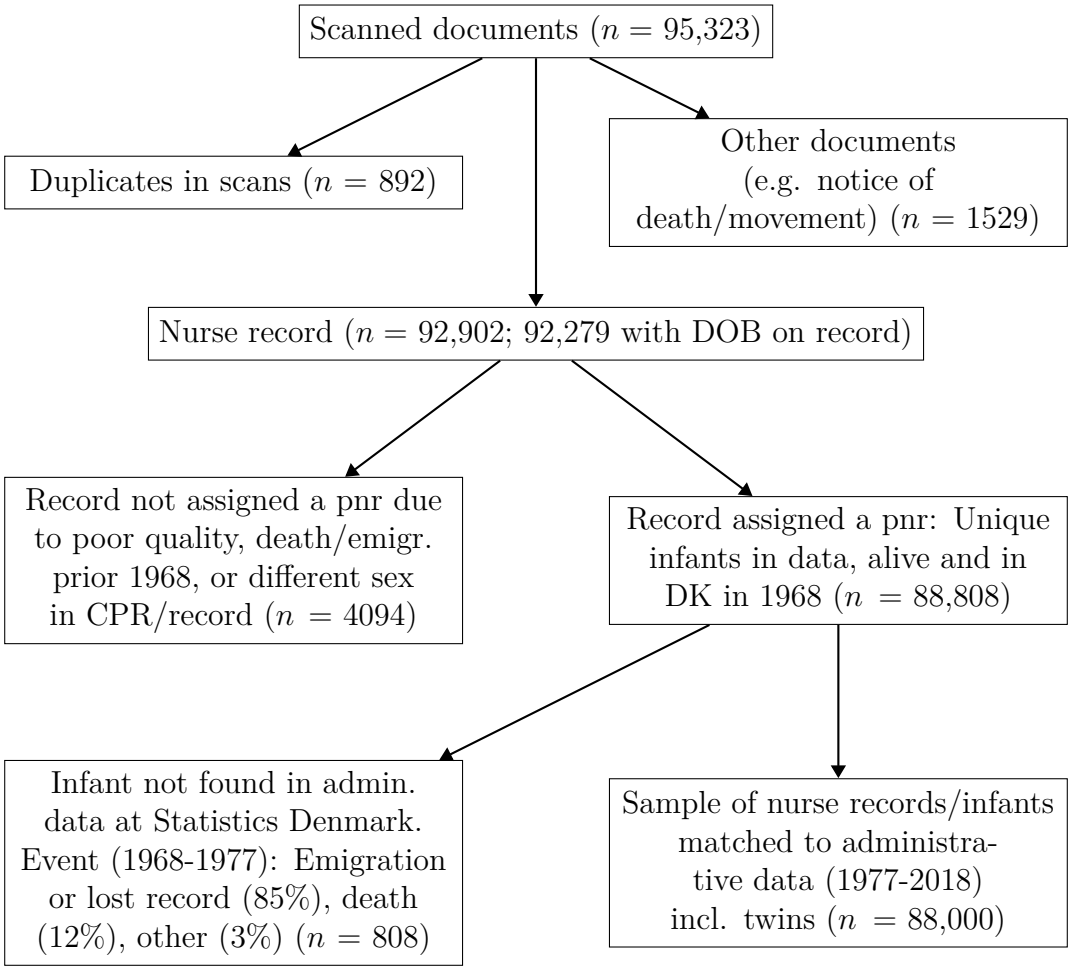


Fig. 2 Flowchart: Merge of Scanned Records to Administrative Data.
Notes: Numbers in parentheses indicate sample sizes.

Was the first-year nurse program in Copenhagen universal and do the preserved (and merged) nurse records represent a good coverage of children in the program? To study these questions, we would optimally like to study individual-level data on Copenhagen residence in the 1960s. Unfortunately, these data do not exist. As we detail in Appendix B, we use aggregate yearly records on the number of infants entering the Copenhagen nurse program to establish that home visiting in Copenhagen was a universal offer for all residents and well-attended, i.e. a vast majority of families entered the program (Copenhagen Statistical Office, various years). Comparing the number of scanned records to aggregate statistics on covered infants in the Copenhagen program, we document that the nurse records likely cover

the universe of (resident) infants entering nurse care in Copenhagen in the period. Finally, the yearly aggregate statistics also document a discontinuation rate of the nurse program in the first year of around 30 percent. This figure implies that up to 30 percent of infants left the program during the first year and thus non-compliance for the second and third year extended nurse treatment is likely to be strongly related to this dropout already prior to assignment of the treatment.[18]

3.2 Administrative Data from Statistics Denmark

We use administrative register data from Statistics Denmark to study measures of human capital, labor market status, and health. We combine data from the education and income registers (1980-2018/2019, respectively), the death register (1970-2020), and the national patient register with discharge diagnoses from hospitals (1977-2018). To be included in our analyses for health and SES outcomes, individuals have to survive until and be observed in the register data at Statistics Denmark after 1977/1980. However, for all individuals in the nurse records with a valid pnr assinged in 1968, we can keep track of pre-1977 outcomes (deaths and emigrations).[19]

To confront the issue of multiple hypothesis testing and to increase statistical power, we create two indices summarizing socioeconomic and health status in adulthood. To form the

[18] Aggregate yearly Copenhagen statistics in Appendix B show that the number of infants covered in the nurse program very closely mirrors the yearly number of observations in our scanned nurse records. For several cohorts, we have more infants in the individual level data than reported in the statistical yearbooks that we use, a pattern that is likely attributable to differences between end of year stocks and flows of infants entering (but also leaving) the program during the year. Unfortunately, no individual level accounts of residence exists in the period that we study. We can, however, use individual level data on place of birth (rather than residence) from Statistics Denmark to show that around two thirds of infants born in Copenhagen in the 1959-1967 period are covered in the scanned nurse records. We also show that Copenhagen births that are not observed in the nurse records are much more likely to be hospital births. At the time, home births were still the norm and hospital births were a clear marker of either health risks and social disadvantage. Hospital births were thus much more likely to not be Copenhagen residents. By further analyzing the birth registrations of all infants covered in the nurse records, and by documenting that a significant share of them (18 percent) were not born in Copenhagen, we provide evidence for the importance of inner-Danish migration early in children's lives. In conclusion, we believe that the nurse records cover the vast majority of nurse program eligible infants (i.e., Copenhagen resident children).

[19] We have obtained information on their latest known status (death, emigration, residence) directly from the CPR register for the 1968-2017 period. Data for the pre-1977 period is typically not available in the research data provided at Statistics Denmark.

SES index, we include information on individuals' years of education, whether the individual obtained a post-secondary education (vocational, short-tertiary, or higher), average earnings during ages 30-50,[20] and the share of time in employment during ages 30-50. In our index of adult health, we aggregate measures on contacts with and diagnoses given in hospitals in the 1977-2018 period. Given the relatively young age of our treated cohorts during the observed outcome period, we summarize measures of inpatient hospital contacts and a set of the diagnoses, that have been related to early-life circumstances. Specifically, we include the number of hospital nights during ages 30-59 and an indicator for ever having been hospitalized, as well as indicators for ever having been diagnosed with one of the following conditions: Diabetes, cardiovascular disease, cancer, asthma, any mental health issues, or infections.[21]

We follow Kling et al. (2007) in creating our indices as equally weighted averages of the z-scores of the underlying variables, orienting the signs such that a more beneficial score is associated with a higher value.[22] We calculate the z-scores by subtracting the mean and dividing by the standard deviation for the control group, i.e., individuals *not* born between the first and the third of a month.[23] We calculate the mean and standard deviation based only on control group individuals for whom we observe all underlying variables of the index. We further rescale our indices by dividing with the control group standard deviations, thereby ensuring unit standard deviation of our indices. This final step has no impact on the statistical significance of our results but ensures we can directly interpret our estimates as differences in standard deviations without rescaling by the standard deviations of our indices.

Importantly, for both our SES and health index (measured over 20 and 29 year periods), all individuals have to survive to at least 1977 and we ignore attrition of individuals (temporarily

[20]We winsorize our earnings measurements (one percent of both tails). We do this after inflation adjusting earnings and separately for each age. Our results are robust to not winsorizing earnings.

[21]For details on the diagnoses included, consult Appendix C. We do not omit birth-related hospital admissions for females. Birth hospitalization lengths have been declining over time. This development should be correlated with being born on the first three days of the months for our estimates to be biased. Importantly, we show no impacts on timing of fertility (any children, number of children and age at first child) for of the extended nurse treatment for focal individuals.

[22]If any one of the underlying variables of the index is not observed for an individual, we drop the individual from the analyses where we use that index.

[23]For summary statistics on the components of the SES and health indices see Table 3.

or permanently) out of the sample during the outcome period if the individual has at least been observed once.[24] Thus selective attrition may bias our findings. To address this issue and to also study mortality and mobility in its own right, we perform two types of analyses: First, we only consider the sub-sample of individuals who survive and live in Denmark in 2017. Second, to directly analyze the impact of day of the month of birth on the probability of ever being observed in our outcome data, of leaving Denmark at any point, and of dying (both in the short and long run), we use all 92,279 records with a manually transcribed date of birth and impute the status of non-merged children from the nurse records as either "dead" or "emigrated". The imputed sample analyses help us assess selection out of our long-run outcome samples (due to emigration and deaths, in particular prior to the administrative data period) but are also interesting in their own right. In a final set of analyses, we study mortality among the individuals that are observed in the mortality register, i.e. the individuals who survive until 1970. As mortality is still relatively low for the age groups that we study (and given that the vast majority of deaths occur early in the cohorts' lives or during the final years of our outcome data), we only perform analyses of coarse mortality outcomes (probability of death) rather than other types of survival analyses.

Finally, we study outcomes of other family members of focal children: For their mothers, we consider fertility (age at birth, spacing between births, and total number of children) and labor market outcomes (average income at ages 40-65 and the share of years the mother is employed during ages 40-65). Theses analyses may shed light on potential mechanisms for impacts on treated children. For the offspring of focal individuals (both male and female children covered in the nurse records), we primarily focus on birth outcomes (birth weight in grams, low birth weight status and preterm status).[25] We also consider educational attainment and young adult earnings for the subset of second generation children born prior

[24]For the SES index, attrition out of the sample leads to measures being missing for the relevant years. The individual will be missing in the index analysis if all years of an underlying SES variable are missing. For the health index, both admissions and diagnoses are set to zero in years where the individual is not observed in the outcome data.

[25]We discuss first generation fertility impacts, which are minor.

to 1991 and thus old enough to be measured in our data. Importantly, this sample of second generation children of appropriate age in our outcome data contains around 34 percent of the full second generation sample.

3.3 The Copenhagen Perinatal Cohort 1959-1961

To explore intermediate childhood impacts of extended nurse care and thus potential mechanisms for longer-run impacts, we use data from the Copenhagen Perinatal Cohort (CPC), a prospective cohort study of initially 9,125 infants born 1959-1961 in Rigshospitalet, the largest hospital in Copenhagen. Mothers were invited for follow-up physician measurements of children in the hospital and completed surveys on children's health and development at ages three and six years (Zachau-Christiansen and Ross, 1975). Additionally, CPC children residing in Copenhagen were examined by school doctors (at ages six, seven, ten and 13 years). A subset of the handwritten data from the CPC follow-up in-hospital measurements, mother surveys, and school doctor measurements have been manually transcribed and used in earlier research (Merrick et al., 1983; Schack-Nielsen et al., 2010). In our analyses, we use these data.

The CPC births are not representative for the Copenhagen population as, at the time, home births were still the norm.[26] Given the large share of non-Copenhagen resident births in the CPC, we constrain our analysis sample to singleton children, who also have a nurse record (using the unique identifier assigned to all children in 1968). In doing so, we likely constrain the sample to Copenhagen residents subject to the trial (N = 4,369). Given attrition and non-response at measurements at ages three and six, we use samples of between 1,800 and 2,500 children in our analyses. While the samples vary in their number of observations across outcomes, attrition should be systematically related to the day of the month of birth for our analyses to be biased. Finally, it deserves mention that we due to data use agreements

[26]Predominantly pregnant women for whom a home delivery was not deemed advisable and socially disadvantaged women (who wanted to conceal their birth and opt for an adoption) were referred to hospitals for their birth. Given an over-representation of children with poor initial health in the CPC, we expect the results of our analyses to reflect the impact of extended nurse visits especially for this group of children.

cannot merge the CPC data set with our transcription of nurse record content, including the treatment table indicator. Thus we perform reduced form analyses using the CPC data and scale those results with the first stage from our main analysis on the full set of nurse records.

We study the following childhood outcomes: First, height is a recognized health measure responsive to childhood health and nutritional investments (Deaton, 2007; Bhalotra and Rawlings, 2013). We consider physician-measured height at ages three, six, seven, ten and 13.[27] Second, we consider a set of mother-reported measures of health and health investments: Indicators for the child having been hospitalized, vaccination compliance, and antibiotics consumption (reported at ages three and six). We measure vaccine uptake as being vaccinated with all vaccinations part of the standard vaccination program at the time (tetanus, polio, pertussis, tuberculosis, and small pox). We measure antibiotics consumption as ever having been treated using antibiotics (penicillin or sulfonamide, primarily reflecting penicillin which was much more commonly used). We interpret our measure of antibiotics consumption as a measure of disease exposure given large-scale and easy access to the relevant drugs in our setting (with around 48 percent of children ever having consumed antibiotics at the given ages). In historical work and work from developing country settings, exposure to early childhood disease has both been linked to height and adult health outcomes (see, for example, Crimmins and Finch, 2006). Third,we use a set of developmental milestone questions ($N = 20$) reporting the mother-assessed age at which the child was able to perform a given task. We group tasks into six domains (developmental milestones with respect to language, motor development, eating, dressing, social interactions, and toilet training) as in Schack-Nielsen et al. (2010) and also create an overall measure summarizing the six milestone domains.

[27] We age-adjust our analyses (as some children are older than others at the time of measurement).

4 Empirical Methods

To identify the effect of the extended nurse program, we exploit the pragmatic trial in Copenhagen and compare average outcomes of children (and families) across birth day of the month groups. In our baseline specification, we regress our outcome Y_i for individual i on Z_i, an indicator for whether individual i was born between the first and the third of a month. Thus we identify the intention-to-treat effect of being born on the first three days of the month. We also provide the related instrumental variable estimates, which can be interpreted as the impact of extended NHV on compliers. We measure the treatment with extended nurse care as the presence of a treatment table in the child's nurse record.

To assess the robustness of our findings (and to demonstrate the validity of our design), we estimate more encompassing specifications: We add cohort fixed effects γ_{cohort} (accounting for potential systematic differences in program implementation across birth cohorts), day of the week fixed effects δ_{dow} (accounting for potential differences in timely access to the nurse program), and family characteristics observed by nurses, X_i. These characteristics include an indicator for the child's low birth weight status (transcribed from the nurse records) and indicators for child sex, whether the child was a firstborn, whether the child was born in Copenhagen, whether the child was born in a hospital, and the mother's age at birth (all observed in the administrative data).

$$Y_i = \alpha + \beta Z_i + \theta' X_i + \gamma_{cohort} + \delta_{dow} + \varepsilon_i$$

Additionally, we consider heterogeneity of the effect of extended care by focusing on subgroups in the population. To begin with, we assess heterogeneity by child sex. Moreover, we consider heterogeneity by other groups defined by characteristics that proxy different dimensions of disadvantage and are easily observed by nurses: Low birth weight status (birth weight below 2500 grams), age of the mother (whether mother was above or below 23–the median in our sample–at the time of birth), birth location (home or hospital), whether we

observe the father in our data, whether the mother has completed more than compulsory education, and breastfeeding status at one month (whether the child is breastfed at all).[28] Finally, given the potential importance of parental experience, we also consider heterogeneity across child parity (whether the child is firstborn or of higher parity).

In all our analyses, we rely on the pragmatic randomization by day of the month for identification and assume that being born between the first and the third day of a month does not have an impact on outcomes of assigned children and families through other channels than the extended nurse program. Further, we rely on the assumption that the outcomes of an individual should not be affected by the assignment of treatment to *other* individuals.

There are two main threats to identification: First, being born between the first and the third of a month could potentially affect the outcomes of children and families through other channels, for example, if families have more resources at the beginning of a month (a payday effect). To asses the potential importance of this factor, we constrain our control group as in a regression discontinuity approach and compare children born on the first three days of a month to those born in the four days after. Moreover, we run our main analyses on a sample of children from the same cohorts but born in the three biggest Danish towns outside Copenhagen: Aarhus, Odense, and Aalborg. As those did not provide extended nurse care assigned in a similar way as in Copenhagen, we do not expect to see any effects of day of the month of birth for these individuals.[29]

Second, the outcome of an individual may be affected by the assignment of additional visits to other individuals. These spillovers, for example within families, may result from parents sharing knowledge with others or asking for counseling also with respect to other children. As families may have more than one child in our sample period, treatment assignment may directly impact fertility decisions and the outcomes of younger siblings. To assess the importance of spillovers, we constrain our sample to a sample of children, who are

[28] We observe mothers' educational attainment in the administrative data after 1980.

[29] In Aarhus, the municipality experimented with longer follow up of at risk families from 1962 onwards. This project, however, selected families according to risk rather than their child's date of birth.

the first observed child of the given family in our data. Finally, we directly study whether mothers change their fertility as a response to being assigned to extended nurse visits.

4.1 Descriptive Statistics

Table 3 presents means and standard deviations for birth, family, and first year background characteristics transcribed from the nurse records (Panel A), background characteristics observed in the administrative data (Panel B), and adult outcomes (Panel C) for the matched sample and subsamples defined by the child's day of the month of birth. There are no systematic differences in background characteristics, first year characteristics and parental inputs (such as nurse-recorded breastfeeding or childcare attendance) across day of the month of birth.[30] For outcome measures, we find no large differences across individual outcomes. However, individuals born between the first and the third of a month appear to be slightly better off in terms of some health measures, in particular hospitalization measures. As apparent in the Table, we consider a young adult population with relatively low disease prevalence. This fact motivates our index construction.

Table 4 presents means and standard deviations for central pre-determined background characteristics (Panel A) and childhood outcomes measures (Panel B) for the full CPC sample and for subsamples, defined by the child's day of the month of birth. Most background characteristics are not systematically different across day of the month of birth, although (significantly) fewer children are born with low birth weight between the first and the third day of the month in our analysis sample. When examining the outcome measures, children born between the first and the third day of the month generally reach development milestones earlier, are taller, are less likely to have used antibiotics, and are more likely to be vaccinated. These larger differences across groups in childhood data may point to immediate (childhood) impacts of the extended nurse treatment, which we more formally assess in the following.

[30] Appendix Table A1 shows the distribution of birth weights in our sample and in the Statistical Yearbook for Copenhagen in 1951-1953 (Copenhagen Statistical Office, 1962 report, Table 171). The distributions are very similar, speaking to the quality of the transcription and the coverage of our data. Birth weights in our sample are slightly right-shifted relative to the aggregate statistics, likely due to us excluding twins.

Table 3 Descriptive Statistics (Means, std.dev.), Main Sample.

Variable	Full sample	Born 4-31	Born 1-3	N
A. Background Characteristics, Nurse Records				
Exclusively breastfed at 1 mo.	0.57 (0.49)	0.57 (0.49)	0.58 (0.49)	72,144
Exclusively breastfed at 2 mo.	0.33 (0.47)	0.33 (0.47)	0.33 (0.47)	69,319
Exclusively breastfed at 3 mo.	0.21 (0.41)	0.21 (0.41)	0.20 (0.40)	67,599
Exclusively breastfed at 6 mo.	0.03 (0.17)	0.03 (0.17)	0.03 (0.17)	65,171
Weight at birth	3326.04 (527.10)	3326.33 (528.47)	3323.09 (512.99)	80,648
Weight at 1 mo.	4078.44 (557.55)	4078.94 (557.07)	4073.38 (562.47)	72,171
Weight at 2 mo.	5043.59 (796.33)	5042.70 (797.82)	5052.62 (781.13)	70,060
Weight at 3 mo.	5876.37 (866.96)	5875.40 (866.59)	5886.24 (870.64)	67,857
Weight at 4 mo.	6682.29 (913.12)	6679.58 (912.27)	6709.43 (921.17)	65,542
Weight at 6 mo.	7953.96 (1002.36)	7952.43 (1002.39)	7969.33 (1002.03)	63,376
Lifts head at 2 mo.	0.48 (0.50)	0.48 (0.50)	0.52 (0.50)	53,392
Lifts head at 3 mo.	0.92 (0.27)	0.92 (0.27)	0.92 (0.27)	56,921
Lifts head at 4 mo.	0.98 (0.13)	0.98 (0.13)	0.98 (0.12)	43,335
Sits at 2 mo.	0.02 (0.13)	0.01 (0.10)	0.07 (0.26)	44,397
Sits at 4 mo.	0.02 (0.15)	0.02 (0.12)	0.08 (0.27)	38,095
Sits at 6 mo.	0.22 (0.41)	0.21 (0.41)	0.27 (0.45)	37,047
Sits at 9 mo.	0.89 (0.31)	0.89 (0.31)	0.91 (0.29)	31,372
Childcare attendance at 6 mo.	0.08 (0.27)	0.08 (0.27)	0.10 (0.30)	55,704
Childcare attendance at 12 mo.	0.07 (0.25)	0.07 (0.25)	0.08 (0.27)	46,456
Good mother mental health at 1 mo.	0.67 (0.47)	0.66 (0.47)	0.68 (0.47)	52,331
Good mother mental health at 12 mo.	0.62 (0.48)	0.63 (0.48)	0.61 (0.49)	48,289
Good mother physical health at 1 mo.	0.76 (0.43)	0.76 (0.43)	0.76 (0.43)	54,182
Good mother physical health at 12 mo.	0.73 (0.44)	0.73 (0.44)	0.73 (0.44)	48,043
All Cells Empty, Nine Months	0.27 (0.45)	0.27 (0.45)	0.26 (0.44)	81,220
Treatment Table	0.06 (0.23)	0.00 (0.05)	0.57 (0.50)	87,336
Any filled	0.05 (0.21)	0.00 (0.05)	0.50 (0.50)	86,632
Share filled	0.04 (0.18)	0.00 (0.03)	0.40 (0.44)	86,632
B. Background Characteristics, Admin. Data				
Female	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	87,336
Born in CPH	0.82 (0.39)	0.82 (0.39)	0.82 (0.39)	86,944
Hosp. birth	0.22 (0.41)	0.22 (0.41)	0.21 (0.41)	87,336
Mother's age at birth	24.98 (5.37)	24.98 (5.37)	24.98 (5.36)	86,754
Firsttime Birth	0.55 (0.50)	0.55 (0.50)	0.55 (0.50)	86,754
Father missing	0.05 (0.23)	0.05 (0.23)	0.06 (0.23)	87,336
First in Cohort	0.74 (0.44)	0.74 (0.44)	0.74 (0.44)	86,754
C. Outcome Measures for Focal Children, Admin. Data				
Yrs. of educ.	13.73 (2.54)	13.73 (2.54)	13.72 (2.55)	85,330
Above compulsory educ.	0.74 (0.44)	0.74 (0.44)	0.74 (0.44)	85,330
Higher education	0.29 (0.45)	0.29 (0.45)	0.29 (0.45)	85,330
Earnings at 25 (DKK 1000)	190.54 (133.81)	190.57 (133.80)	190.30 (133.93)	85,741
Avg. Earnings 30-50 (DKK 1000)	286.50 (181.80)	286.31 (181.92)	288.27 (180.71)	85,035
Avg. Empl. age 30-50	0.80 (0.29)	0.80 (0.29)	0.81 (0.28)	84,100
Out of Labor Force, Age 50	0.17 (0.38)	0.17 (0.38)	0.16 (0.37)	78,595
Diabetes	0.04 (0.20)	0.04 (0.20)	0.04 (0.19)	87,336
Cardiovascular disease	0.27 (0.44)	0.27 (0.44)	0.26 (0.44)	87,336
Heart disease	0.05 (0.22)	0.05 (0.23)	0.05 (0.22)	87,336
Asthma	0.05 (0.22)	0.05 (0.22)	0.05 (0.21)	87,336
Cancer	0.07 (0.25)	0.07 (0.26)	0.07 (0.25)	87,336
Any mental health contact	0.14 (0.35)	0.14 (0.35)	0.14 (0.35)	87,336
Infection	0.19 (0.39)	0.19 (0.39)	0.19 (0.39)	87,336
Hosp. nights 20-29	8.07 (19.71)	8.08 (19.72)	8.01 (19.65)	87,336
Hosp. nights 30-39	6.12 (19.45)	6.14 (19.63)	5.92 (17.70)	87,336
Hosp. nights 40-49	4.77 (19.20)	4.81 (19.15)	4.47 (19.65)	87,336
Hosp. nights 50-59	3.61 (16.12)	3.62 (16.04)	3.49 (16.85)	87,336
Ever hospitalized (30-59)	0.76 (0.43)	0.76 (0.43)	0.75 (0.43)	87,336

Notes: Income is measured in 1000DKK (2015). The diagnoses measure whether an individual has ever been diagnosed with the condition in our outcome data.

Table 4 Descriptive Statistics (Means, std.dev.), CPC Sample.

Variable	Full sample	Born 4-31	Born 1-3	N
A. Background Characteristics, CPC				
Female	0.49 (0.50)	0.49 (0.50)	0.43 (0.50)	2,704
Social status (0=low, 1=middle, 2=high)	1.15 (0.82)	1.15 (0.82)	1.13 (0.83)	2,314
Smoking in last trimester of pregnancy	0.53 (0.50)	0.54 (0.50)	0.52 (0.50)	2,655
Preterm Birth	0.17 (0.37)	0.17 (0.38)	0.14 (0.35)	2,223
Birth order	1.76 (1.12)	1.75 (1.12)	1.83 (1.10)	2,704
Birth weight (g)	3214.09 (576.46)	3209.45 (583.62)	3257.69 (503.08)	2,704
Low birth weight (<2500g)	0.09 (0.29)	**0.10 (0.29)**	**0.06 (0.23)**	2,704
B. Outcome Measures, CPC				
Height (cm), age 3	95.89 (4.06)	95.81 (4.03)	96.59 (4.22)	1,206
Height (cm), age 6	118.55 (5.48)	118.43 (5.43)	119.56 (5.88)	1,090
Height (cm), age 7	122.92 (5.31)	122.86 (5.33)	123.39 (5.10)	1,773
Height (cm), age 10	138.15 (6.35)	138.13 (6.31)	138.27 (6.72)	2,400
Height (cm), age 13	155.43 (7.91)	155.48 (7.91)	154.95 (7.98)	2,348
Number of vaccinations (3 years)	2.41 (0.90)	2.40 (0.90)	2.42 (0.95)	2,402
Number of vaccinations (6 years)	2.05 (1.03)	2.05 (1.03)	2.09 (1.02)	2,716
Fully Vaccinated at 3 Years	0.60 (0.49)	0.60 (0.49)	0.65 (0.48)	2,402
Fully Vaccinated at 6 Years	0.44 (0.50)	0.44 (0.50)	0.46 (0.50)	2,716
Ever consumption of antibiotics at 3 years	0.48 (0.50)	0.48 (0.50)	0.39 (0.49)	2,273
Ever consumption of antibiotics at 6 years	0.47 (0.50)	0.48 (0.50)	0.34 (0.47)	1,888
Ever Hospital admission, age 1-3	0.27 (0.44)	0.27 (0.44)	0.25 (0.43)	2,402
Ever Hospital admission, age 3-6	0.32 (0.47)	0.33 (0.47)	0.29 (0.46)	2,716
Otitis, bronchitis or pneumonia, age 1-3	0.30 (0.46)	0.30 (0.46)	0.26 (0.44)	2,704
Otitis, bronchitis or pneumonia, age 3-6	0.31 (0.46)	0.31 (0.46)	0.28 (0.45)	2,743
Mean Milestone z-score: Language	-0.00 (1.00)	0.01 (1.00)	-0.05 (1.01)	2,001
Mean Milestone z-score: Motor	0.00 (1.00)	0.01 (1.00)	-0.09 (0.98)	2,180
Mean Milestone z-score: Eating	-0.00 (1.00)	0.01 (1.00)	-0.14 (1.00)	1,946
Mean Milestone z-score: Dressing	0.00 (1.00)	-0.00 (1.00)	0.02 (1.03)	1,698
Mean Milestone z-score: Soc. Inter.	-0.00 (1.00)	0.02 (0.99)	-0.16 (1.07)	2,021
Mean Milestone z-score: Toilet	0.00 (1.00)	0.00 (1.00)	-0.02 (1.02)	2,134
Mean Milestone z-score: Combined	-0.00 (1.00)	0.01 (1.00)	-0.11 (0.99)	2,256

Notes: All summary statistics are based on the CPC sample of children, who have a valid registrations on the relevant outcome measure. For background characteristics we report descriptives of children, who were participating at least in the three year examination. The height measures beyond age six come from school doctor measurement. Bold figures indicate significant differences at five percent level.

5 Results

5.1 First Stage: The Assignment of Extended Nurse Visiting

Figure 3 shows the share of matched nurse records that we classify as containing a treatment table across the days of the month (our first stage). Around 57 percent of children covered in the nurse records and born between the first and the third of a month have this type of record. Among those not born on these days this figure is 0.3 percent. Across birth months, quarters, and years, we observe a stable share of records classified as treatment records, indicating that

the pragmatic trial was in place in the entire data period. We also find that the share of treatment table records is very similar across geographic units defined by parishes within the municipality of Copenhagen (see Appendix Table A2). This finding supports that the trial covered all of Copenhagen.

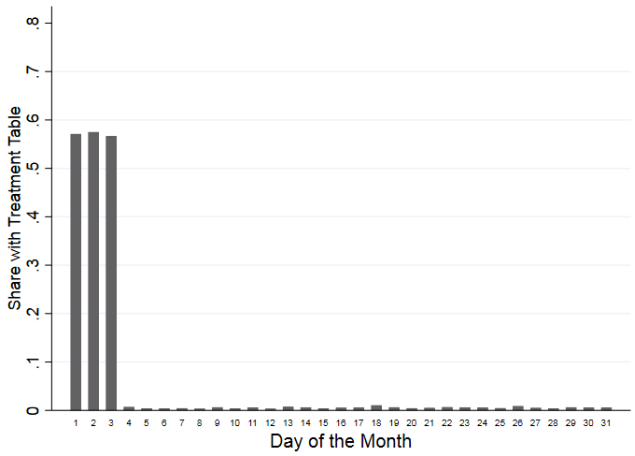
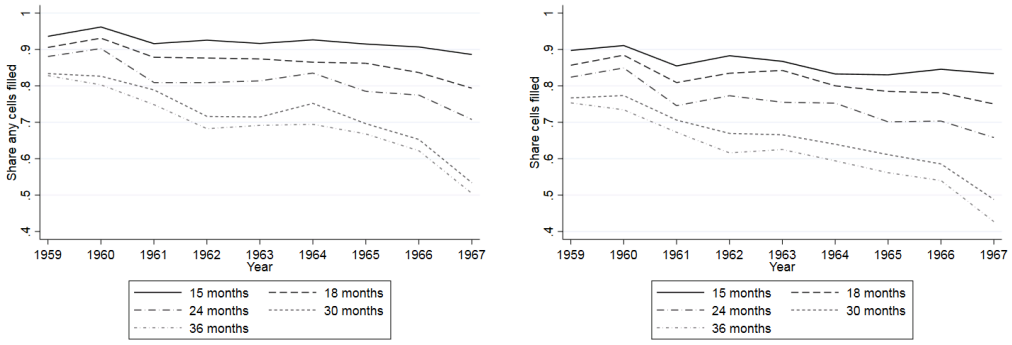


Fig. 3 Share of Records with a Treatment Table by Day of the Month of the Child's Birth.

The treatment table in a child's record implies that the treatment was formally assigned. Figure 4 illustrates the intensity of the treatment that was administered. We use two related measures based on our classification of cell entry for each designated cell (at visits at ages 15, 18, 24, 30, and 36 months): Whether at least one cell was filled out and the share of cells that were filled out.[31] Between 50 and 95 percent of the treatment table records have at least one cell filled out, with the highest share of filled cells at the visits during the second year of the child's life. Although the records are not completely filled out, these findings indicate that the presence of a treatment table suggests that the visits were in fact performed. For the final years in our sample period, there is a tendency for fewer registrations in the child's third year of life. Focusing on the average number of assigned visits, we find that children born on the first three days of the month receive 2.6 and 1.4 visits of the scheduled three and two visits in the second and third year of life, respectively.

[31] We calculate the share as the number of cells that were filled out divided by the total number of relevant cells. If a treatment table is not present, the share is set to zero.



(a) Any cell filled **(b)** Share of eligible cells filled

Fig. 4 Completeness of Nurse Registrations in Treatment Tables Across Visits and Birth Cohorts.

Notes: Panel 4a shows the share of treatment table records where at least one cell is filled out and Panel 4b shows the share of cells filled out (by year of birth of the child).

Table 5 presents regression equivalents of the graphical evidence on the first stage. The first stage holds across specifications, successively adding relevant fixed effect and pre-treatment control variables and when restricting the sample to only those born between the first and the seventh of a month. In all specifications, being born between the first and the third of a month increases the probability of having the treatment table with around 57 percent.[32]

To further probe the robustness of our first stage, Appendix Table A4 confirms that the first stage also holds using all nurse records (including nurse records not merged to outcome data and thus not included in the main analyses).[33] Moreover, Appendix Table A5 shows that the first stage is also robust to selecting nurse records that we more confidently classify as Copenhagen residents at child age one year (around the time of treatment assignment): We condition our sample to only include records that have at least one nurse registration at 12 months (identified in our transcription of records). Additionally, we add a more extensive

[32]Appendix Table A3 shows the equivalent estimates for the probability of at least one cell being filled out, and the share of cells being filled out (with all unfilled cells in all records being set to zero).

[33]Moreover, Appendix Figure A2 graphically illustrates that the share of treatment table records is clearly different across the day of the month groups even when we include nurse records that cannot be matched to outcome data: For individuals who survive to 1968 and thus have an identifier, we find a very similar pattern of treatment table records across the days of the month. For those without an assigned identifier (dead or outside Denmark prior to 1968), we see a much lower share of treatment tables. These records are both more likely to be of poor quality (both complicating our data transcription and merge) and likely to include a large share of infants who either died or moved out of Denmark during their first year of life. Thus they likely contain a larger share of families who did not receive the treatment offer at child eligibility age one.

set of controls for factors likely observed by nurses prior to treatment assignment to our list of controls (indicators for a good economic status of the family, for good maternal mental and physical health, for an orderly home, and for exclusive breastfeeding, all recorded at child age one month). Using this sample and our set of main specifications, we find a strong and slightly larger first stage of around 0.73-0.75. This finding is in line with more complete records belonging to children, who remain residents in Copenhagen and thus have a higher probability of being included in the trial at one year of life. The sample is, however, also selected on our ability to transcribe the relevant records and therefore we prefer our full sample of merged nurse records for the main analyses.

Table 5 First Stage: Presence of a Treatment Table.

	(1)	(2)	(3)	(4)	(5)
Treatment Table Present					
Born 1-3	0.568***	0.568***	0.568***	0.567***	0.566***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
No. of obs.	87,336	87,336	87,336	79,806	17,770
F-value	11,330	1,267	761	444	450
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: Each cell shows estimates from a separate regression. The controls added in columns (4) and (5) include maternal age at birth and indicators for child sex, child low birth weight status, the child being born in Copenhagen, the child being born in a hospital, and the father being observed in the administrative data. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Finally, given our data on child and family characteristics measured prior to the initial nurse contact, we also formally assess whether the first stage differs across subgroups in a complier analysis. This analysis may also reveal if nurses targeted certain types of children and families in the trial (even though this was explicitly not intended). As we illustrate in Appendix Table A6 (using our first stage specification with cohort and day of the week fixed effects), the day of the month of birth predicts the presence of a treatment table across subgroups in our sample defined by relevant characteristics that nurses observe during

their early contacts with the families. We observe a strong first stage across groups, which suggests a generally good compliance with universal assignment. One aspect that deserves mention is that children born at hospitals and children with a missing father registration appear to be less likely to be assigned to extended nurse care (there is considerable overlap in the two groups). We attribute this fact to the special pattern in hospital births at the time: Socially disadvantaged mothers (including non-Copenhagen residents) and mothers with health issues were more likely to give birth at hospitals while home birth remained the norm for healthy mothers until the 1970s. Thus we think that larger mobility and also larger first-year mortality (and thus drop-out out of the nurse program prior to treatment assignment) explains this pattern.

5.2 Main Results

Long-run Impacts of Extended Home Visiting on Adult Socioeconomic and Health Outcomes Table 6 presents reduced form and instrumental variable estimates for our two main outcomes, the SES and health indices. Mirroring the descriptive statistics, there is no significant effect of birth group on the SES index and estimates are very small. The reduced form results suggest a significant positive effect on our health index of 0.026-0.028 standard deviations. Our instrumental variable results show a significant increase of 0.044-0.05 standard deviations.

As Table 6 illustrates, our main results are robust across specifications. Pointing to the successful implementation of the pragmatic randomization, the treatment and control groups are similar across a host of observable characteristics and their inclusion as controls does not impact our estimates. We further probe the robustness of the main estimates in two ways: First, we perfom the main analysis focusing on the first child of each family that is observed in the trial period. In this analysis we show that also in a sample that makes spillovers less likely, our results are very similar (Appendix Table A7).[34] Second, as illustrated in

[34] An alternative way of alleviating concerns about spurious effects is a placebo permutation test where we randomly assign the day of the month of birth to individuals in our sample. We calculate t-values associated

Table 6 Main results: The effect of extended nurse care on SES and health indices.

	(1)	(2)	(3)	(4)	(5)
Panel A: SES index – ITT					
Born 1-3	0.007	0.006	0.006	0.005	0.008
	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)
No. of obs.	83,029	83,029	83,029	76,257	16,973
Panel B: SES index – LATE					
Treatment Table	0.012	0.010	0.010	0.007	0.012
	(0.020)	(0.020)	(0.020)	(0.022)	(0.027)
No. of obs.	83,029	83,029	83,029	76,257	16,973
Panel C: Health index – ITT					
Born 1-3	0.028***	0.027**	0.026**	0.028**	0.026*
	(0.011)	(0.011)	(0.011)	(0.012)	(0.015)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel D: Health index – LATE					
Treatment Table	0.050**	0.047**	0.047**	0.048**	0.044*
	(0.019)	(0.019)	(0.019)	(0.021)	(0.026)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5. * p<0.1 ** p<0.05 *** p<0.01.

Appendix Table A8, there is no similar impact of being born in the first three days of the month in Aarhus, Odense and Aalborg, the three largest Danish towns outside Copenhagen. An analysis of data from children in these towns for the 1959-1967 cohorts directly assesses potential day of the month of birth effects (in an alternative way to only using a subset of children born in the first week of each month).[35]

with regressing our indices on an indicator for a birth in the first three days of a month, repeating this exercise 1000 times. This approach is inspired by MacKinnon and Webb (2016) and conducted similarly to Sievertsen and Wüst (2017). Appendix Figure A3 plots the cumulative distributions of estimates for the impact of random day of the month on outcomes illustrating that our result for a long run health effect–but not effect on the SES index–falls in the tail of this distribution.

[35]To construct this sample we select individuals of the 1959-1967 cohorts with a birth parish in one of the three towns. Thus we cannot identify residency in this sample. Further, we do not control for birth weight in our specification adding pre-treatment controls, as we do not observe this variable in the administrative data. Instead, we add a town fixed effect to our list of controls.

Our main index results conceal some heterogeneity across underlying outcome measures and samples of treated children. First, separately analyzing underlying measures of health, educational attainment, and labor market outcomes included in the indices, we find that estimates for education are economically small and insignificant across different margins of the education distribution (Appendix Table A10).[36] For employment measures, we detect small positive impacts (significant at the 10 percent level) for the average share of the time individuals are in employment (ages 30-50). Related to this finding, we see a small and not significantly estimated negative impact on the probability of being out of the labor force (i.e., most likely on disability pension) at age 50, an age at which we observe all individuals in our outcome data and an age where a non-trivial share of the population starts exiting the labor force (17 percent).[37] Given the relatively young age of the individuals in our sample, our finding for the health index may bear relevance in the longer run also for economic outcomes, such as these retirement decisions often related to health issues. Studying the set of underlying health variables in the health index, we find positive impacts on health across the board (i.e., negative estimates for ever being diagnosed) but most estimates lack statistical significance individually. As exceptions, diagnoses for asthma and cancer stand out (Appendix Tables A11 and A12). We have in supplementary analyses confirmed that our main results for the health index are robust to the omission of both asthma diagnoses and asthma-related hospital admissions (Appendix Table A13).

Second, we find that the long-run impacts are larger for treated girls.[38] While we not

[36]To illustrate that our results are not dues to a lack of statistical power, we conduct a simple power analysis for SES outcomes. Our tests report the required ITT effect on outcomes to find an effect with 80 percent power and at five percent significance. Appendix Table A9 shows that the required effect sizes in our setting are all relatively small to moderate. For example, an effect on average income at ages 25-50 would have to be no greater than 5207 2015 DKK, i.e. an increase of around two percent of the mean of the dependent variable. In contrast, Hoehn-Velasco (2021) shows that county-level health departments in rural areas of the US rolled-out during 1908–1933 increased male earnings by up to five percent for those exposed before the age of five. This program covered between 40–100 percent of the population, offering around two examinations per infant.

[37]Our SES index results are sensitive to the inclusion of this final variable with its inclusion resulting in small positive and significant estimates.

[38]Compliance with treatment assignment was similar across subgroups in our population and thus selective assignment of treatment should not drive these findings (Appendix Table A6).

formally can reject equality across estimates, this conclusion holds for the health index results and results for the impact of the trial on the employment-related components of the SES index (employment share and the probability of leaving the labor force early), which are stronger for girls (with significant impacts for the mentioned SES outcomes across specification, as shown in Appendix Table A14 for females and Appendix Table A15 for males). Moreover, when we divide our SES index in two indices of educational and labor market measures, respectively, we find significant impacts of girls (ITT impact of 0.004 standard deviations) but not boys on the labor market index (Appendix Table A16).

Considering heterogeneity across other dimensions—initial conditions with respect to children's health, parental SES and investment decisions in the first year—Table 7 presents our results for relevant subgroups. While we generally cannot reject equality of estimates across subgroups, we find large health effects for low birth weight children (which are statistically different from the estimates in the no-low birth weight group). For low birth weight children, our ITT (0.141) and instrumental variable estimates (0.251) document large impacts on the health index. Relating these estimates to the mean of the outcomes variable in the control group of low birth weight children (-0.20) illustrates that this is an economically meaningful impact. It compensates treated low birth weight children for the relative health disadvantage that they face.

Table 7 The Effect of Extended Nurse Care on the SES and Helath Indices across Subgroups, Reduced Form and Instrumental Variable Estimates.

	(8) Female		(1) Low BW		(2) Firstborn		(3) Mother $\leq$ 23		(4) Hosp. birth		(5) Miss. father		(6) Mother $\geq$ comp. edu.		(7) Breastfed	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
A. SES Index:																
Born 1-3	-0.010	0.022	0.005	0.041	0.008	0.001	-0.006	0.012	0.003	0.006	0.007	0.004	0.011	-0.005	0.029	-0.011
	(0.017)	(0.015)	(0.013)	(0.056)	(0.017)	(0.016)	(0.018)	(0.015)	(0.013)	(0.026)	(0.012)	(0.053)	(0.015)	(0.017)	(0.020)	(0.017)
P-value (Difference)	0.156		0.526		0.786		0.428		0.929		0.956		0.467		0.129	
Treatment Table	-0.018	0.039	0.008	0.073	0.012	0.003	-0.012	0.020	0.005	0.012	0.012	0.008	0.019	-0.010	0.049	-0.018
	(0.030)	(0.027)	(0.023)	(0.100)	(0.028)	(0.029)	(0.035)	(0.025)	(0.021)	(0.055)	(0.021)	(0.107)	(0.026)	(0.030)	(0.034)	(0.028)
MDV	0.06	-0.02	0.04	-0.25	-0.08	0.11	-0.08	0.09	0.09	-0.20	0.04	-0.27	-0.20	0.29	-0.09	0.14
No. of obs.	42,118	40,911	72,874	3,781	36,902	45,698	32,004	50,596	65,006	18,023	79,027	4,002	42,394	37,489	29,221	39,500
B. Health Index:																
Born 1-3	0.023	0.033**	0.021*	0.141**	0.035**	0.020	0.013	0.035**	0.023*	0.036	0.030***	-0.033	0.025	0.027*	0.040**	0.029*
	(0.016)	(0.015)	(0.012)	(0.059)	(0.017)	(0.015)	(0.018)	(0.014)	(0.012)	(0.025)	(0.011)	(0.049)	(0.016)	(0.016)	(0.019)	(0.017)
P-value (Difference)	0.644		0.047		0.521		0.324		0.629		0.210		0.942		0.669	
Treatment Table	0.041	0.058**	0.036*	0.251**	0.057**	0.038	0.025	0.059**	0.039*	0.076	0.052***	-0.067	0.043	0.049*	0.067**	0.047*
	(0.028)	(0.027)	(0.022)	(0.107)	(0.027)	(0.028)	(0.035)	(0.023)	(0.021)	(0.052)	(0.020)	(0.099)	(0.027)	(0.028)	(0.033)	(0.027)
MDV	0.11	-0.12	0.01	-0.20	-0.03	0.03	-0.03	0.02	0.03	-0.10	-0.00	0.01	-0.08	0.08	-0.03	0.04
No. of obs.	44,646	42,690	76,631	4,017	38,868	47,886	33,590	53,164	68,237	19,099	82,591	4,745	44,342	39,061	30,820	41,324

Notes: Each cell shows estimates from separate regressions estimated across subgroups of the population (denoted in the column headings). Regressions include fixed effects but exclude control variables. Breastfed equals one if the child was breasted at one month of age (registered by the nurse). MDV is the mean of the dependent variable in the control group for the relevant subgroup, i.e. those not born between the first and the third of a month. The p-value is for a test of equality of coefficients from a fully interacted model. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Considering only the robustly established health impacts, we can conclude that the extended home visiting program was highly cost effective: Abstracting from fixed costs for office facilities and management, nurses' yearly wages and transportation allowances amounted to 47,000 DKK/357,541 DKK (1970/2020) (DNBH, 1970). Their responsibility for approximately 160 children per nurse and year in Copenhagen suggests an estimated per-child cost of only around 293 DKK/2,234 DKK (amounting to 342 USD (2020)).[39] As described in Section 2, the extended program for 10 percent of families was introduced within the budgetary frame of the existing one year program (by implementing a one-two visit decrease in first year visits for all families in Copenhagen and thus in our sample). Thus the documented long-run health benefits—especially among disadvantaged children—are likely to have outweighed the costs of the program (which are admittedly a lower bound estimate) to a great extend.

Short- and Long-run Mortality and Emigration Outcomes A factor ignored in our analyses thus far is the impact of extended nurse care on mortality and emigration (both of which lead to individuals not being observed in our outcome data in either all or some years): First, non-merged nurse records may indicate early-life mortality or emigration. If extended nurse care increased the survival (and thus merge with outcome data) of weaker children, we may find attenuated average impacts on survivors' outcomes. Second, for the individuals initially merged to outcome data, the outcomes that we generate ignore that some individuals leave the data in the data period (or are missing in single years).

We perform two sets of analyses: One that includes non-matched individuals and imputes outcomes, and one that focuses on long-run survivors (i.e. individuals for whom we are certain that they are alive and in Denmark by 2017). First, Table 8 examines mortality and emigration events as outcomes. Our measure for emigrations is equal to one if the individual has ever left Denmark (and was registered as an emigrant). Thus it is not conditional on never returning to Denmark, while mortality is an absorbing state.

[39]The Statistical yearbook for Copenhagen in 1963 reports a total of 56 nurses and a total of 8,879 infants under their supervision (Copenhagen Statistical Office, 1963).

We start by examining whether we are more likely to observe the treatment group in our outcome data. Columns one of Table 8 suggests that being born on the first three days of the month increases this probability with 0,4 percent. In columns 2-7 we analyze emigration and mortality outcomes. For both types of outcomes, we either impute non-merged individuals with a one or a zero or only use merged individuals (our main analysis sample).[40]

Turning to ever emigrating from Denmark, we find no differences across treated and control children. For the probability of being dead by 2020, the final year of our mortality data, our ITT estimates suggest that children born on the first three days of the month are less likely to die (a seven percent change in our merged analysis sample). Imputing non-merged individuals as being dead (early-life mortality) increases this estimate. Thus our findings may imply that children born on the first three days of the month had a higher probability of surviving childhood and also in the long run experience (emerging) mortality advantages.[41] This finding in turn would mean that "weaker children" survived longer due to the trial and thus the positive adult health effects are a sign for health improvements outweighing this negative health selection.

In a second analysis (further probing the robustness of our main results to potentially selective attrition out of our sample), we re-estimate our main analyses for the SES and health index on a sample of survivors (and Danish residents) observed in the year 2017. This essentially eliminates all individuals who either die or (permanently) emigrate from Denmark. Our results for this sample confirm the conclusions from above: We find very similar estimates for the SES index and, if anything, larger effects for the health index in this sample of survivors (Appendix Table A19). Thus in their combination these results suggest that we, if anything, underestimate the health benefits of the treatment.[42]

[40]The majority of non-merged individuals are not merged to the 1968 CPR registry and thus have no personal identifier.

[41]In supplementary analyses we have truncated our mortality data at the year 2016 and 2016, respectively. We show that point estimates are similar in these samples but mortality results are significant likely due to lower mortality in the shorter follow-up window.

[42]In analyses on our sample of individuals from the three othermajor Danish towns, we do not find differential impacts of day of the month of birth on mortality (1977-2017) or the probability of ever leaving Denmark (1977-2014), see Appendix Table A20. As we only observe individuals from the towns outside

Table 8 The Effect of Extended Nurse Care on Mortality and Ever Leaving Denmark, Reduced Form Estimates.

	Ever in data	Ever Emi. imp. (1)	Ever Emi. imp. (0)	Ever Emi. obs	Death imp. (1)	Death imp. (0)	Deaths obs.
Born 1-3	0.004**	-0.007	-0.003	-0.004	-0.010***	-0.006**	-0.006**
	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
MDV	0.961	0.190	0.151	0.158	0.119	0.080	0.083
No. of obs.	92,279	92,279	92,279	88,751	92,279	92,279	88,751

Notes: All regressions are based on a specification including year of birth and day of the week of birth fixed effects. The sample sizes in columns (1)-(3) and (5)-(6) are larger than for our main analyses as we include all nurse records, i.e. also individuals, who are not merged with the administrative data. Deaths and emigration (including temporary leaving Denmark) are observed for all identified nurse record individuals in the 1968-2020 period. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Mechanisms: Childhood Health Impacts for Focal Children and Impacts on Mothers of Focal Children To better understand the drivers of longer-run impacts, we use complementary data from a cohort study to examine immediate childhood impacts of the extended nurse treatment. As the CPC data only cover a non-representative subset of Copenhagen (hospital) births at the time, we may detect (larger) childhood health impacts among the relatively disadvantaged children, who participate at the CPC followup data collections.[43]

Figure 5 shows the intention-to-treat results (alongside 95 percent confidence intervals) for birth on the first three days of the month on childhood height, mother-reported vaccination completion and antibiotics usage, mother-reported infectious disease prevalence and hospital admission events, and mother-reported ages at completion of developmental milestones. Appendix Figure A4 plots estimates from separate regressions for girls and boys. Appendix Tables A21 through A24 present the respective estimates and sample sizes across outcomes (combined samples).[44]

Copenhagen (and thereby without preserved nurse records) if they survive to 1977 and are observed only in our administrative data from Statistics Denmark, the definition of mortality and emigration is slightly different to our main analysis (i.e. we do not observe individuals who die or emigrate prior to the onset of the register data at Statistics Denmark in the 1968-1977 period.

[43]Our CPC analyses are purely reduced form analyses, as we cannot merge the CPC samples with the information transcribed from the nurse records due to data use agreements (as detailed in Section 3). We scale all results with the implied first stage from our main analyses.

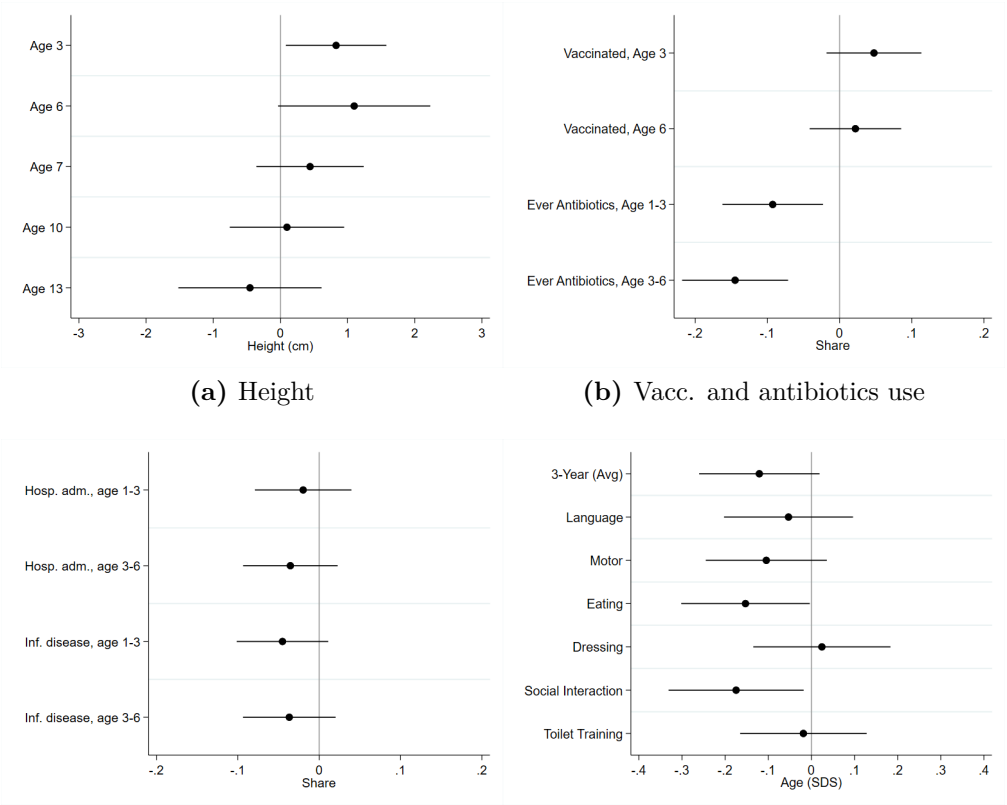
[44]In these tables we both present estimates including and excluding a low birth weight control, given

The figures illustrate three main patterns: First, children born on the first three days of the month are taller during early childhood but those impacts fade out over time. As Appendix Figure A4 illustrates, this pattern is driven by boys. Second, while there is no impact on birth of the first three days of the month on full uptake of the suggested childhood vaccinations (our primary measure of uptake of preventive care), we find that children born on those days are less likely to ever have consumed antibiotics during early childhood (both measured for ages 1-3 and 3-6). While not statistically significant at conventional levels in the full sample, negative estimates for the probability of having been exposed to infectious diseases (primarily pneumonia and otitis) and the probability of having been hospitalized support this finding. Third, there is some suggestive evidence for children born on the first three days of the month being younger at completion of developmental milestones. However, also most of those estimates are imprecise.

Our estimates for antibiotics usage are large in magnitude: Children born the first three days of a month are 14.5 percentage points less likely to have received antibiotics at age six, which is around 30 percent less likely than those born any other day of the month (around 48 percent ever received antibiotics at age six). Scaling the results by the implied first stage, we find that treated children see a 25 percentage point decrease in antibiotics use. The finding that treated children were dramatically less likely to consume antibiotics during childhood resonates with our result that in the longer-run, children in the treated group have a lower probability of being diagnosed with asthma, a disease that has been related to childhood antibiotics use (Patrick et al., 2020). While combined estimates for hospital admissions and infectious disease are not significant at conventional levels, we find that this finding conceals some heterogeneity across girls and boys with somewhat larger and some significant health benefits for girls (as illustrated in Appendix Figure A4). This finding is similar to our long-run health result.

In their combination our results indicate that extended nurse visits primarily were pro-

evidence for baseline imbalance in this characteristic in the CPC samples that we use.



(a) Height **(b)** Vacc. and antibiotics use

(c) Hosp. adm. and inf. disease **(d)** Development milestones

Fig. 5 Short-run Effects of Extended Nurse Care on Health, Health Care Use, and Development Milestones, ITT.

Notes: The figures present estimates for the impact of being born on the first three days of the month on outcomes (ITT), as well as 95 percent confidence intervals, for children observed in both the nurse records and the CPC follow-up. See Section 3.3 for further details.

ductive in promoting childhood health. However, the available milestone measures may not be adequate to capture other dimensions of children's development (such as cognitive ability). The dosage and content of the extended nurse treatment, alongside other expanding programs such as universal childcare, may explain why we see very limited impacts of extended preventive care on measures child development and later of educational attainment. Similarly, the finding of some fade out of health effects (measured as height) may suggest that later investments can alleviate initial disadvantages, in line with earlier work on the interaction of targeted preschool and nurse home visit exposure in Denmark (Rossin-Slater and Wüst, 2020). Given the likely exposure of all children in our sample to various follow-

up interventions such as universal preschool and health checks during schooling years this substitution likely plays a role also for long-run impacts.

Given the focus of the extended nurse program on families' well-being and the timing of visits during a period where families make various decisions about aspects such as fertility and labor supply, another relevant mechanism for long-run impacts is program impacts on other family members, in particular mothers. If present, fertility or labor supply effects on mothers may be a pathway for impacts on focal children. Appendix Table A25 presents analyses for maternal outcomes, focusing on mothers who gave birth to their first child in the 1959-1967 period (and thus are represented in our data with their first and potentially also later children). We do not observe economically or statistically significant differences in fertility decisions or labor market outcomes for these first-time mothers. We conclude that it is unlikely that family adjustments at the margin of fertility or maternal labor supply drive our findings for health outcomes of focal children.

First Generation Fertility and Impacts on the Next Generation Do the positive health impacts on treated children in the long run extend to focal children's offspring? To factor in benefits beyond focal children, we turn to an analysis of second generation outcomes in a last step. Health at birth is a natural starting point for this analysis in the light of extensive research documenting strong inter-generational ties in measures such as low birth weight (Currie and Moretti, 2007) and given its predictive power for later life outcomes. Inter-generational links in health at birth likely operate both through biological processes defined as early as in utero and through other channels. Those channels may include adult socioeconomic status, partner choice, health and life styles. Thus we may expect that health investments in toddlers, which impact their own health trajectory, may benefit their offspring through these channels.

Because we find larger long-run health effects of extended nurse visits for first generation individuals with an initial health disadvantage, we also focus on studying second generation impacts in a "high impact sample", namely the offspring of first generation individuals with

initial health disadvantages. Importantly, as we run into small cell problems given the share of treated children and the share of children with low birth weight at birth among them (our main measure of poor health at birth), we define a broader sample of focal children with low initial health. We proxy low health for the focal child if at least one of the following conditions holds: the focal child was low birth weight, had a young mother (below age 21, i.e. among the 25 percent youngest mothers), was born in a hospital or had a missing father registration. In our main second generation analyses, we consider all observed children born to the focal mothers and fathers to further increase sample size but also present results for first parity children (parity is always defined as biological parity, i.e. birth order for a given mother). For the second generation children, who have both parents observed in our trial data (because focal children partnered, which is the case for around 20 percent of second generation children), we keep the mother spell to create a unique sample of second generation children. In all second generation analyses, we control for focal child year and day of the week of birth fixed effects as well as focal child sex.

Before we turn to second generation results, we study potential impacts on focal children's fertility decisions (the margin of having children at all, the number of children conditional on having a child and the age at first birth). Overall, we find little indication for fertility impacts for focal children (Appendix Table A26). When considering heterogeneity, we show that conditional on having any children, treated males tend to have slightly fewer children overall, an effect that is economically small (an ITT effect of 0.039 children, an at a mean of 2.1 children, see Appendix Table A27). We do not find fertility responses in our poor initial health sample (Appendix Table A28).

Table 9 shows our results for the impact of first generation treatment exposure on second generation health at birth outcomes. We divide the sample of offspring by the low health status measure defined for the first generation.[45] While most estimates are very imprecise, we find suggestive evidence for the offspring of focal children with low initial health status

[45]Results for average impacts on the second generation are very imprecise and do not allow for strong conclusions (Appendix Table A29).

facing a decreased risk of being low birth weight. Effects for second generation children are qualitatively similar for a sample focusing on first parity second generation children and are also similar but very imprecise across the smaller samples focusing only on offspring of focal mothers or focal fathers, respectively (Appendix Tables A30). These findings carefully suggest that the large health improvements, which we observe for focal children with an initial health disadvantage, may spill over to the health of the second generation.

Table 9 The Effect of Extended Nurse Care and Birth Outcomes in the Second Generation (ITT). Heterogeneity by Focal Child Initial Health

Second Gen.	Birth Weight (g)		Low BW		Preterm	
First Gen. Low Health	0	1	0	1	0	1
A. All Second Gen. Children						
Born 1-3	-8.241	13.860	0.000	-0.008**	0.004	-0.002
	(8.012)	(9.798)	(0.003)	(0.004)	(0.003)	(0.004)
MDV	3451.90	3382.88	0.05	0.07	0.06	0.07
No. of obs.	67,162	46,301	67,162	46,301	66,320	45,611
B. First Parity Second Gen. Children						
Born 1-3	-9.808	17.349	-0.001	-0.014**	0.003	-0.002
	(11.110)	(13.658)	(0.005)	(0.006)	(0.005)	(0.006)
MDV	3367.06	3299.43	0.06	0.08	0.07	0.07
No. of obs.	33,203	22,584	33,203	22,584	32,664	22,117

Notes: Each cell shows estimates from a separate regression (for second generation birth outcomes) across subgroups of the population (defined by the initial health status of the focal individual, who is now a parent). The unit of observation is the second generation child. The samples consist of all children born to focal individuals (panel A) or first parity children (panel B). If a child has both parents (focal children) in our data, we keep only one child spell and assign treatment and low health status of the mother to the second generation child. The estimates come from our heterogeneity specification including fixed effects for focal children's (now parents') year of birth and day of the week of birth, as well as focal child sex. MDV is the mean of the dependent variable in the control group for the relevant subgroup. * p<0.1 ** p<0.05 *** p<0.01.

Finally, constraining our analysis to the around 34 percent of second generation children born prior to 1991 and thus all have completed their education in our outcome data period, Appendix Table A31 considers educational attainment and young adulthood earnings. On average we find no second generation impacts. For focal children in the good initial health group, we find some significant negative impacts on offspring outcomes of being born on the first three days of the month (on the number of years of schooling and the probability of

having completed more than compulsory education). These results are somewhat sensitive to our selection of second generation cohorts included and economically small for compulsory educational attainment.[46] As such, our results for second generation outcomes remain less conclusive and for measures of poor health at birth constrained by power issues.

5.3 Discussion of Effect Sizes and Relation to Literature

Are our findings for adult health effects of toddler preventive care large (at around 0.05 standard deviations for treated children for our good health index at ages 30-59)? To put our results in context, we compare them to those obtained for other childhood health policies with similar components. In general, those studies find larger long-run health (and socioeconomic) impacts. This finding is likely due to the relevant counterfactual and the intensity and target groups of the program under consideration: As discussed from the onset of this paper, the vast majority of (RCT and observational) empirical work for now has considered high-intensity programs for disadvantaged families, a margin with scope for large impacts.

Hoynes et al. (2016) find that early-life exposure to food stamps in the 1960s to 1980s had significant and large impacts on adulthood metabolic health. Full childhood exposure (in utero to age five) results in a 0.3 standard deviation reduction of metabolic syndrome in a high impact sample (ITT) of poor families likely very dependent on nutritional support. Scaled by the implied first stage, this effect translates into around 0.7 standard deviations for treated children. While the food stamp treatment is much more intensive than what we study in this paper and the average family in our sample was not likely to be exposed to food insecurity, there was scope for improved quality of toddler nutrition encouraged by nurses, especially among disadvantaged families. In line with the findings in Hoynes et al. (2016), we find larger long-run health impacts for children who face initial disadvantages and health-related labor market returns for girls. In follow-up work on the food stamp program and its

[46] In contrast, also in constrained (older) samples of second generation children, we find significant positive impacts of birth in the first three days of the month on health at birth for offspring of focal children with poor initial health (for measures of birth weight and low birth weight).

long-run returns, Bailey et al. (2020a) do not find impacts on a set of (limited) health-related measures (disability reported in survey data), but document benefits on human capital, self-sufficiency and adult survival of food stamp exposure during childhood. Again, exposure during the zero to five year period had the largest returns (a 11 percent decrease in mortality for treated children). Our mortality estimates for the impact of extended nurse care are of similar magnitude and thus arguably large (a decrease in mortality on treated children of around 12.7 percent at a baseline mean of 8.3 percent).

Importantly, nurse visits were not only about nutritional advice but a composite treatment also encouraging other parental investments and uptake of additional health care if deemed necessary. In this respect, our estimates may be comparable to work focusing on the long-run effects of childhood exposure to health insurance (and thus better access to more and better preventive care). Again those studies primarily come from the US and focus on targeted rather than universal interventions. Wherry et al. (2018) and Miller and Wherry (2019) study the impacts in young adulthood of exposure to medicaid expansion around birth and in the first years of life. Wherry et al. (2018) find lower hospitalization rates at age 25 for medicaid exposed children, especially the ones from poor areas and especially related to chronic illness. Miller and Wherry (2019) show that in utero and early-life exposure to medicaid decreases chronic conditions around ages 19-36 by around 1.1 standard deviation for treated children. They also document decreases in hospitalizations due to diabetes and obesity. These estimates are well in line with our (much smaller) estimates for related health measures especially in our sample of children with initial health disadvantage.

Closely related to our work on toddler universal care, earlier work has studied the extensive margin of program roll out of first year universal well-infant care in Denmark, Norway and Sweden. These studies have documented health and mortality benefits of the initial introduction of those first-year programs for the 1930s and 1940s cohorts (living in a high-infant mortality setting). In Denmark, the 1930s and 1940s introduction of nurse visiting decreased the probability of being diagnosed with cardiovascular disease by 1.3-2.8 percent at a control

mean of 26.6 percent in the age group 45-64 years. In Norway, the estimated impact of infant care introduction on a bad health index measured at age 40 was large at 0.19-0.29 standard deviations (Hjort et al., 2017; Bütikofer et al., 2019). Both of these studies focus on first year preventive care (with a counterfactual of no or very little access to regular preventive care for infants).[47] Our work on follow-up care (set in the 1960s in a lower-mortality setting) demonstrates smaller but traceable average long-run health benefits for younger adults and along other dimensions of adult health (including conditions such as asthma, which has been related to early childhood experiences and infections).

Finally, work on intergenerational impacts of early life interventions is still scarce—mainly due to large data requirements. Existing studies have examined targeted interventions using large-scale population data: prenatal exposure to medicaid (East et al., 2022) and exposure to targeted early education programs/preschools (Rossin-Slater and Wüst, 2020; Barr et al., 2022). While we view our intergenerational results as more suggestive, our estimates for measures of health at birth are similar to results in East et al. (2022). They document relatively large impacts on the risk of being low birth weight in the second generation of medicaid exposure of first generation women (2.6 percentage points treatment on the treated impact at a mean of seven percent in treated states). We find a 1.4 percentage points decrease for treated second generation children in the probability of being low birth weight in our "high impact sample" (focusing on first generation children with initial health disadvantage). Thus our findings carefully support this work on the relevance of childhood health interventions—which in our case do not capture the role of biological processes in utero and at birth—for the health in the second generation.

While much of the work referred to here documents effects of the studied interventions beyond health (including measures of educational attainment, labor market outcomes, neighborhood quality, or crime), our results for extended nurse care are concentrated in the health

[47] Closely related is also work from the similarly timed expansion of public health care centers in rural US areas, documenting labor market benefits for exposed boys likely due to improved long-run health rather than higher educational attainment of treated cohorts (Hoehn-Velasco, 2021).

domain (rather than education) and indicate some spill-overs to measures of employment and labor market exit decisions in the longer-run. Several factors may explain this pattern: First, our analyses on childhood data suggest that the health components of the program were particularly effective, a pattern that is likely important also for longer-run health. We show childhood impacts along dimensions that are sensitive to both health information, counseling, and parental health investments, among them early childhood height (a measure strongly related to cumulative inputs in health and nutrition) and disease such as infections. Second, dosage and nurse qualifications may play important roles in explaining the effectiveness of the program: Even though nurses also were supposed to encourage adequate parent-child interactions and advised parents on issues such as child social interactions, our findings suggest that these components were not strong enough (dosage) or that their specific content was not effective in affecting educational attainment. Importantly, nurses had extensive training in the health domain but less experience in toddler development in other domains. Third, even if there were initial program impacts in other domains than health, such as cognition, those may fade out over time, as documented in other early life programs such as early education (Barnett, 1995; Karoly et al., 1998). Related, children of the given cohorts were exposed to multiple programs in the expanding Danish welfare state during childhood, including universal preschool (that witnessed a rapid and massive expansion in the 1960s alongside large-scale maternal entrance on the labor market) and school doctor checkups, which may have served as partial substitutes and thus attenuate our findings in the education domain. Moreover, children grew up in a setting with strongly expanding educational opportunities. In the end, health-focused nurse visits as an "add on" to these major policy and societal changes may not impact outcomes such as education in this context with universal access, even in the disadvantaged groups that we single out.

6 Discussion and Conclusion

While causal evidence across settings documents that early-life health shocks matter for short- and long-run outcomes of exposed children, much less research has been conducted so far on the long-run and potentially multigenerational impacts of positive interventions, in particular on universal policies in place in many settings. The design of these early-life investments in young children's health and development is likely to crucially matter for their returns.

We contribute novel evidence on this topic: Studying a large-scale government trial in Copenhagen with extended universal nurse home visiting for toddlers of the 1959-1967 cohorts, we document long-run benefits of childhood preventive care for the adult health of treated children but no impact on a combined index of adult educational attainment and labor market outcomes. Related to the positive health impacts in mid adulthood, we find some positive impacts of the treatment in terms of employment during adulthood. We find that especially treated women are less likely to leave the labor market at age 50, an age where individuals at an increasing rate start entering programs such as disability pension (subject to health testing). These labor market impacts may become more important as these individuals age. Finally, as the cohorts enter age ranges with higher mortality risks, we find emerging positive impacts on survival. Taken together, these results document large health returns of (early-life) preventive care.

While our focal individuals may still be too young to show the full impact of the program, we find that the health benefits of the program are particularly large in disadvantaged children. We show that children with an initial health disadvantage see health benefits that compensate for this disadvantage and that may even spill over to the health of their offspring. This finding is important, as extended nurse care was not assigned to individuals based on measures of early disadvantage. Still we find the strongest impacts of the universally accessible treatment on disadvantaged groups defined by characteristics that are easily observed by nurses.

Our findings suggest that targeting of a toddler health program like the one studied here–and potentially giving nurses a role to play in the assignment of families–could further improve cost effectiveness. Having said that, the adjustment of the program to cover second and third year visits for some children came at very modest costs. Given its implementation for around 10 percent of the population of children in the trial, it is fair to assume that it could have been implemented to a similarly large group of children selected on measures of disadvantages at the given low cost.

Besides our conclusions on the impacts of extended nurse visiting, our study breaks new ground in making handwritten data on early life circumstances available for research. Thus it exemplifies the great potential for research that transcribed historical records bring into the reach of applied researchers. The analyses of this paper have become possible due to automated and thus faster (and cheaper) transcription of handwritten and structured data. We believe that these types of transcribed data in combination with other, administrative data sources will open up for exiting avenues for future research on important topics such as inter-generational mobility in health and economic outcomes in the Nordic countries and beyond.

References

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of political Economy*, 114(4):672–712.

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

Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in sweden. *The Quarterly journal of economics*, 124(4):1729–1772.

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.

Bailey, M. J., Hoynes, H. W., Rossin-Slater, M., and Walker, R. (2020a). Is the social safety net a long-term investment? large-scale evidence from the food stamps program. Technical report, National Bureau of Economic Research.

Bailey, M. J., Timpe, B. D., and Sun, S. (2020b). Prep school for poor kids: The long-run impacts of Head Start on human capital and economic self-sufficiency. Working Paper 28268, National Bureau of Economic Research.

Barker, D. J. P. (1990). The fetal and infant origins of adult disease. *BMJ: British Medical Journal*, 301(6761):1111.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 5(3):25–50.

Barr, A., Eggleston, J., and Smith, A. A. (2022). Investing in Infants: the Lasting Effects of Cash Transfers to New Families*. *The Quarterly Journal of Economics*. qjac023.

Belfield, C. R., Nores, M., Barnett, S., and Schweinhart, L. (2006). The high/scope perry preschool program cost–benefit analysis using data from the age-40 followup. *Journal of Human resources*, 41(1):162–190.

Besl, P. and McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.

Bhalotra, S., Karlsson, M., and Nilsson, T. (2017). Infant health and longevity: Evidence from a historical intervention in Sweden. *Journal of the European Economic Association*, 15(5):1101–1157.

Bhalotra, S. and Rawlings, S. (2013). Gradients of the Intergenerational Transmission of Health in Developing Countries. *The Review of Economics and Statistics*, 95(2):660–672.

Bluche, T., Louradour, J., and Messina, R. (2017). Scan, attend and read: End-to-end handwritten paragraph recognition with MDLSTM attention. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 1050–1055. IEEE.

Bluche, T., Ney, H., and Kermorvant, C. (2014). A comparison of sequence-trained deep neural networks and recurrent neural networks optical modeling for handwriting recognition. In Besacier, L., Dediu, A.-H., and Martin-Vide, C., editors, *Statistical Language and Speech Processing*, pages 199–210, Cham. Springer International Publishing.

Burbank, A. J., Sood, A. K., Kesic, M. J., Peden, D. B., and Hernandez, M. L. (2017). Environmental determinants of allergy and asthma in early life. *Journal of Allergy and Clinical Immunology*, 140(1):1–12.

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

Clinchant, S., Déjean, H., Meunier, J., Lang, E. M., and Kleber, F. (2018). Comparing machine learning approaches for table recognition in historical register books. In *2018*

13th IAPR International Workshop on Document Analysis Systems (DAS), pages 133–138. IEEE.

Copenhagen Archives (nd). various documents.

Copenhagen Statistical Office (various years). *Statistical Yearbooks*. Bianco Lunos Bogtrykkeri A-S, Copenhagen.

Crimmins, E. M. and Finch, C. E. (2006). Infection, inflammation, height, and longevity. *Proceedings of the National Academy of Sciences*, 103(2):498–503.

Cubuk, E. D., Zoph, B., Shlens, J., and Le, Q. V. (2020). RandAugment: Practical automated data augmentation with a reduced search space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 702–703.

Currie, J. and Almond, D. (2011). Human capital development before age five. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4B, pages 1315–1486. Elsevier, North-Holland, Amsterdam.

Currie, J. and Moretti, E. (2007). Biology as destiny? short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2):231–264.

Dahl, C. M., Johansen, T. S. D., Sørensen, E. N., Westermann, C. E., and Wittrock, S. F. (2021). Applications of machine learning in document digitisation. *arXiv:2102.03239*.

Dahl, C. M., Johansen, T. S. D., Sørensen, E. N., and Wittrock, S. F. (2022). HANA: A HAndwritten NAme database for offline handwritten text recognition. *arXiv:2101.10862*.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences*, 104(33):13232–13237.

Deng, J., Dong, W., Socher, R., Li, L., Li, K., and Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. IEEE.

Det Statistiske Departement (1964). *Spædbørnsdødelighed 1931-1960 [Infant Mortality 1931-1960]*. Det Statistiske Departement.

DNBH (1970). Betænkning Nr 573: Sundhedsplejerske Institutionen [Report on the Danish Home Visiting Program]. Technical report, The Danish National Board of Health.

East, C. N., Miller, S., Page, M., and Wherry, L. R. (2022). Multi-generational impacts of childhood access to the safety net: Early life exposure to medicaid and the next generation's health. Technical report.

Ester, M., Kriegel, H. P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, pages 226–231. AAAI Press.

Forsdahl, A. (1979). Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease? *British Journal of Preventive and Social Medicine*, 31(2):91–95.

Gern, J. E., Lemanske Jr., R. F., and Busse, W. W. (1999). Early life origins of asthma. *The Journal of Clinical Investigation*, 104(7):837–843.

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., and Shet, V. (2013). Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv:1312.6082*.

Goodman-Bacon, A. (2018). Public insurance and mortality: evidence from medicaid implementation. *Journal of Political Economy*, 126(1):216–262.

Harris, C. G. and Stephens, M. (1988). A combined corner and edge detector. In *Proceedings of the Alvey Vision Conference*, pages 147–151.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition.

In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778.

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

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

Hjort, J., Sølvsten, M., and Wüst, M. (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.

Hoehn-Velasco, L. (2021). The long-term impact of preventative public health programs. *The Economic Journal*, 131(634):797–826.

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–34.

Karoly, L. A., Greenwood, P. W., Everingham, S. M. S., Hoube, J., Kilburn, M. R., Rydell, C. P., Sanders, M., and Chiesa, J. (1998). *Investing in Our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions*. RAND Corporation, Santa Monica, CA.

Kim, I. (2020). pytorch-randaugment. `https://github.com/ildoonet/pytorch-randaugment`.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kreiner, C. T. and Sievertsen, H. H. (2020). Neonatal health of parents and cognitive development of children. *Journal of health economics*, 69:102247.

Lee, C. and Osindero, S. (2016). Recursive recurrent nets with attention modeling for OCR in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2231–2239.

MacKinnon, J. G. and Webb, M. D. (2016). Difference-in-differences inference with few treated clusters. Queen's Economics Department Working Paper 1355, Kingston (Ontario).

Merrick, J., Teasdale, T., and Merrick, Y. (1983). School health screening of a birth cohort: a prospective longitudinal study. *International journal of rehabilitation research. Internationale Zeitschrift fur Rehabilitationsforschung. Revue internationale de recherches de readaptation*, 6(4):461—468.

Miller, S. and Wherry, L. R. (2019). The long-term effects of early life medicaid coverage. *Journal of Human Resources*, 54(3):785–824.

Myronenko, A. and Song, X. (2010). Point set registration: Coherent Point Drift. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 32, pages 2262–2275. IEEE.

Noghanibehambari, H. (2022). Intergenerational health effects of medicaid. *Economics Human Biology*, 45:101114.

Olds, D. L., Henderson, C. R., Chamberlin, R., and Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics*, 78(1):65–78.

Olds, D. L., Henderson, C. R., Cole, R., Eckenrode, J., Kitzmann, Harriet Luckey, D., Pettitt, L., Sidora, K., Morris, P., and Powers, J. (1998). Long-term effects of nurse home visitation on childrens criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *JAMA*, 280(14):1238–1244.

Olds, D. L., Kitzman, H., Anson, E., Smith, J. A., Knudtson, M. D., Miller, T., Cole, R., Hopfer, C., and Conti, G. (2019). Prenatal and infancy nurse home visiting effects on mothers: 18-year follow-up of a randomized trial. *Pediatrics*, 144(6).

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32, pages 8026–8037.

Patrick, D. M., Sbihi, H., Dai, D. L. Y., Al Mamun, A., Rasali, D., Rose, C., Marra, F., Boutin, R. C. T., Petersen, C., Stiemsma, L. T., Winsor, G. L., Brinkman, F. S. L., Kozyrskyj, A. L., Azad, M. B., Becker, A. B., Mandhane, P. J., Moraes, T. J., Sears, M. R., Subbarao, P., Finlay, B. B., and Turvey, S. E. (2020). Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: Evidence from population-based and prospective cohort studies. *The Lancet Respiratory Medicine*, 8(11):1094–1105.

Rossin-Slater, M. and Wüst, M. (2020). What is the added value of preschool for poor children? Long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics*, 12(3):255–286.

Schack-Nielsen, L., Sørensen, T. I. A., Mortensen, E. L., and Michaelsen, K. F. (2010). Late introduction of complementary feeding, rather than duration of breastfeeding, may protect against adult overweight. *The American Journal of Clinical Nutrition*, 91(3):619–627.

Schulz, L. C. (2010). The dutch hunger winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences*, 107(39):16757–16758.

Shen, Z., Zhang, R., Dell, M., Lee, B. C. G., Carlson, J., and Li, W. (2021). LayoutParser: A unified toolkit for deep learning based document image analysis. *arXiv:2103.15348*.

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

Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*.

Sutskever, I., Martens, J., Dahl, G., and Hinton, G. (2013). On the importance of initialization and momentum in deep learning. In *International Conference on Machine Learning*, volume 30, pages 1139–1147.

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer Science & Business Media.

Van den Berg, G. J., Deeg, D. J., Lindeboom, M., and Portrait, F. (2010). The role of early-life conditions in the cognitive decline due to adverse events later in life. *The Economic Journal*, 120(548):F411–F428.

van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(11):2579–2605.

Wherry, L. R., Miller, S., Kaestner, R., and Meyer, B. D. (2018). Childhood medicaid coverage and later-life health care utilization. *Review of Economics and Statistics*, 100(2):287–302.

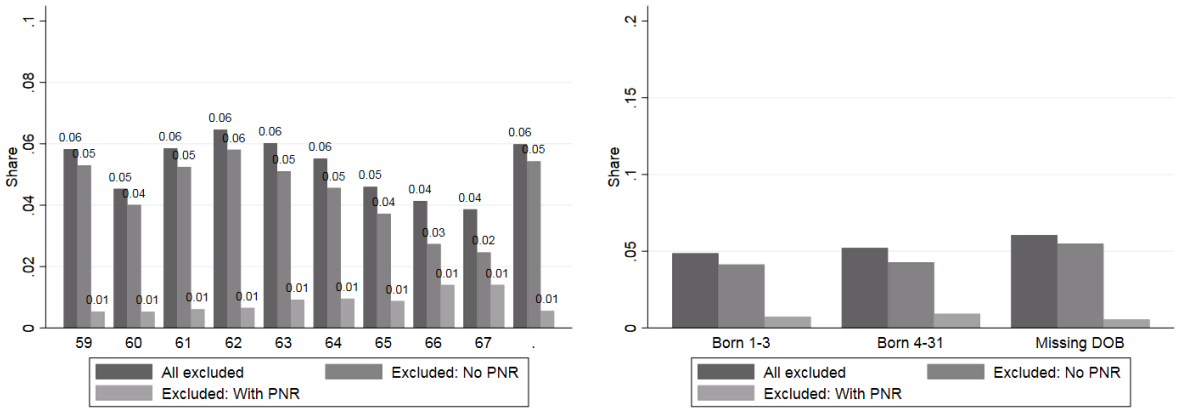
Wüst, M. (2012). Early interventions and infant health: Evidence from the Danish home visiting program. *Labour Economics*, 19(4):484–495.

Xie, S., Girshick, R., Dollár, P., Tu, Z., and He, K. (2017). Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1492–1500.

Zachau-Christiansen, B. and Ross, E. M. (1975). *Babies: Human Development During the First Year*. John Wiley & Sons, New York, NY.

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

A Additional Tables and Figures



(a) Excluded by year **(b)** Excluded by day of the month (groups)

Fig. A1 Share of Excluded Nurse Records (with and without a unique personal identifier, pnr), by Birth Cohort and by Day of the Month of Birth, 1959-1967.

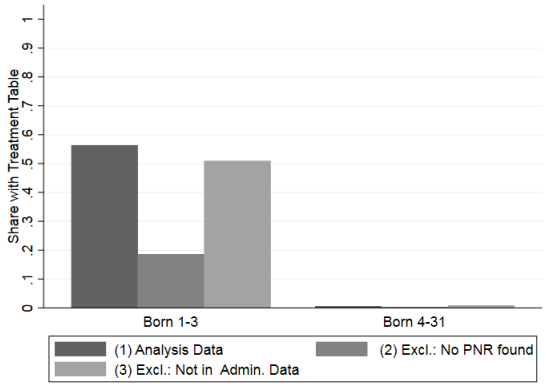
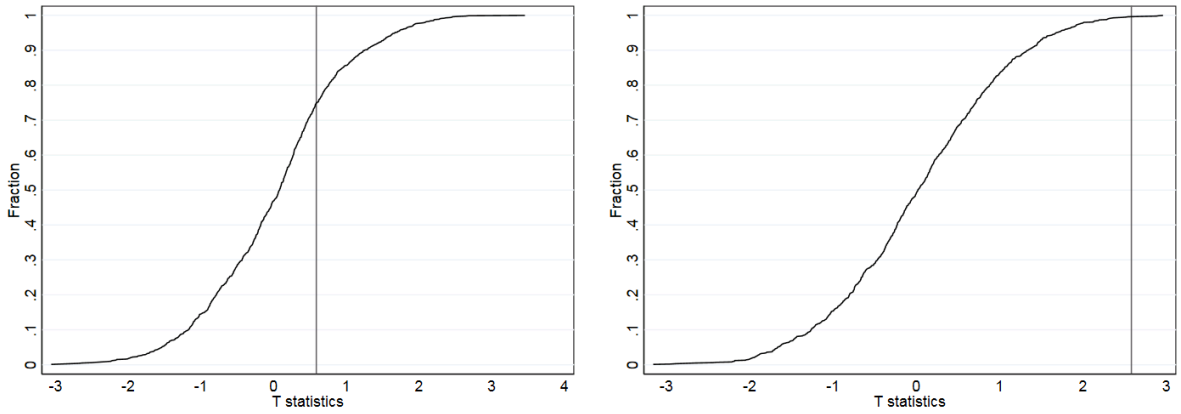


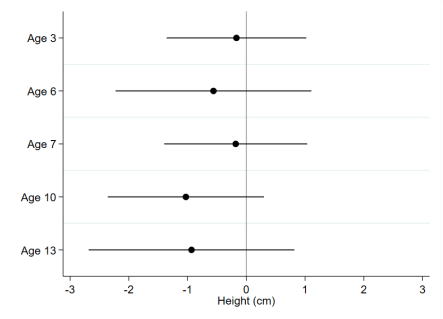
Fig. A2 Share of Records with a Treatment Table by Day of the Month of Birth. Matched and Unmatched Nurse Records.



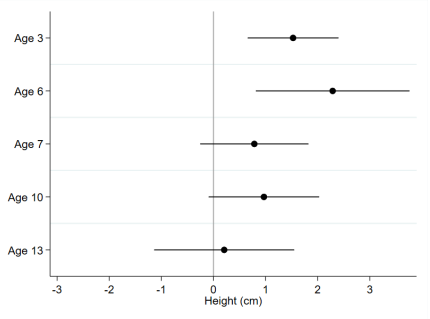
(a) SES index **(b)** Health index

Fig. A3 Placebo permutation test: t-values from Random Allocation of Day of the Month of Birth.

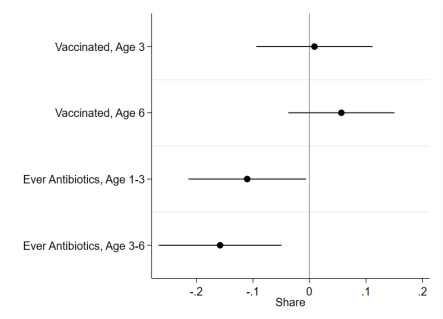
Notes: The figure shows the empirical cumulative distribution function of placebo t-values from 1000 regressions of the outcome on whether an individual is born the first three days of the month. We assign the day of the month of birth at random. The vertical line shows the t-value of the true estimate. Panel (a) shows the results for our SES index and Panel (b) shows the results for our health index.



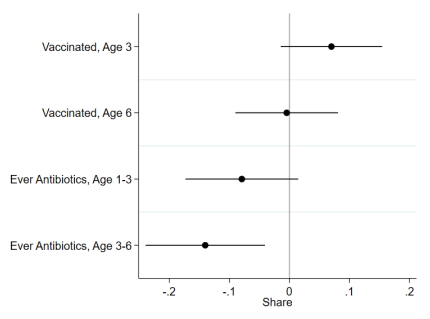
(a) Height (cm), female



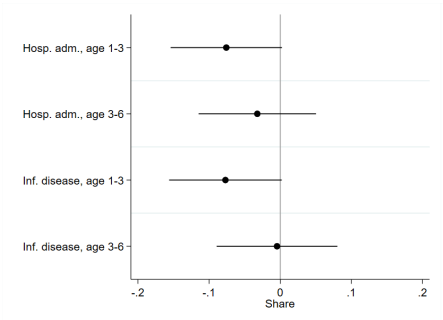
(b) Height (cm), male



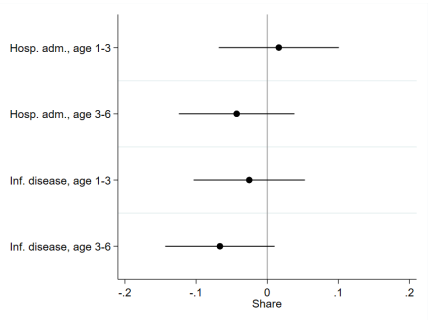
(c) Vacc. and Antibiotics, female



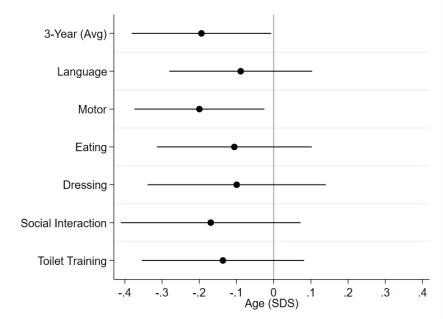
(d) Vacc. and Antibiotics, male



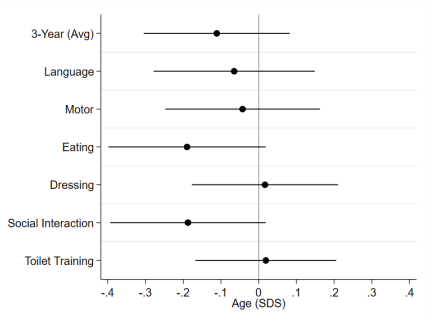
(e) Hosp. adm. and inf., female



(f) Hosp. adm. and inf., male



(g) Development milestones, female



(h) Development milestones, male

Fig. A4 ITT Effects on Childhood Health, Health Care Use and Development Milestones.

Notes: The figures present estimates (and 95pct. confidence intervals) for the impact of being born on the first three days of the month on child outcomes. The samples are children observed in both the nurse records and the CPC follow-up. See Section 3.3 for further details.

Table A1 Birth Weight in the Analysis Sample and in the Copenhagen Statistical Yearbook for the Years 1951-1953.

	Male			Female		
Birth weight	No. obs.	Share (%)	Share YB (%)	No. obs.	Share (%)	Share YB (%)
Below 1500	113	0.3	0.2	126	0.3	0.4
1500-2000	371	0.9	1.1	438	1.1	1.3
2000-2500	1,421	3.4	4.7	1,718	4.3	6.2
2500-3000	5,845	14.1	19.3	7,700	19.4	25.0
3000-3500	14,926	36.0	37.9	16,151	40.6	40.3
3500-4000	13,333	32.1	27.6	10,622	26.7	21.6
4000-4500	4,671	11.3	8.0	2,617	6.6	4.7
4500-5000	713	1.7	1.2	332	0.8	0.5
At or above 5000	106	0.3	0.0	37	0.1	0.0

Notes: The table shows the distribution of birth weight in the analysis sample, i.e. the number of singleton children (m/f) from the records matched to administrative data and the share of individuals in the records in the given birth weight group. The final column for both boys and girls denotes the share of the respective birth weight group in the aggregate Copenhagen Statistical Yearbooks (YB).

Table A2 Share of Treatment Table Records among all Records Across Areas in Copenhagen.

Treatment table	No		Yes	
	No. of obs.	Share (%)	No. of obs.	Share (%)
Bispebjerg Provstri	2,544	93.46	178	6.54
Holmens Provsti	2,308	93.52	160	6.48
Nordvestre Provsti	1,450	93.31	104	6.69
Nordøstre Provsti	6,121	93.34	437	6.66
Søndre Provsti	10,829	93.41	764	6.59
Vestre Provsti	4,148	94.90	223	5.10
Vor Frue Provsti	21,083	93.56	1,452	6.44
Hospital birth	18,145	95.00	954	5.00
Other	15,216	94.84	828	5.16

Notes: The table shows the number and share (%) of individuals without and with an identified treatment table in their record across different *provsti* (geographical units consisting of several parishes) in our sample. The *other* category are individuals with unknown birth registration codes, religious groups with own birth registration codes, or birth registration codes outside Copenhagen due to, e.g., migration into Copenhagen shortly after birth.

Table A3 First Stage: Presence of a Treatment Table and Completeness of Registrations. Main Analysis Sample.

	(1)	(2)	(3)	(4)	(5)
Panel A: Treatment Table Present					
Born 1-3	0.568***	0.568***	0.568***	0.567***	0.566***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
No. of obs.	87,336	87,336	87,336	79,806	17,770
F-value	11,330	1,267	761	444	450
Panel B: Any Cell of Treatment Table Filled					
Born 1-3	0.502***	0.502***	0.502***	0.504***	0.503***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
No. of obs.	86,632	86,632	86,632	79,230	17,196
F-value	7,987	898	539	314	317
Panel C: Share of Cells of Treatment Table Filled					
Born 1-3	0.394***	0.394***	0.394***	0.397***	0.396***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	86,632	86,632	86,632	79,230	17,196
F-value	6,409	719	431	250	253
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: Each cell shows estimates from a separate regression. The controls added in columns (4) and (5) include maternal age at birth and indicators for child sex, child low birth weight status, the child being born in Copenhagen, the child being born in a hospital, and the father being observed in the administrative data. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A4 First Stage: Presence of a Treatment Table and Completeness of Registrations. All Nurse Records, including Non-Matched Records

	(1)	(2)	(3)	(4)
Panel A: Treatment Table Present				
Born 1-3	0.552***	0.552***	0.552***	0.552***
	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	92,279	92,279	92,279	21,285
F-value	11,186	1,253	752	746
Panel B: Any Cell of Treatment Table Filled				
Born 1-3	0.487***	0.486***	0.486***	0.487***
	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	91,545	91,545	91,545	20,564
F-value	7,913	894	537	530
Panel C: Share of Cells of Treatment Table Filled				
Born 1-3	0.381***	0.381***	0.381***	0.382***
	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	91,545	91,545	91,545	20,564
F-value	6,368	719	431	427
Cohort FE	No	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes
Only born 1-7	No	No	No	Yes

Notes: Each cell shows estimates from a separate regression. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01. For further details, see Notes for Table A3.

Table A5 First Stage: Presence of a Treatment Table, Sample of Children with Non-empty Nurse Records at One Year.

	(1)	(2)	(3)	(4)	(5)
Born 1-3	0.729***	0.729***	0.728***	0.757***	0.757***
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)
No. of obs.	57,897	57,897	57,897	34,107	7,541
F-value	14,849	1,690	1,017	390	395
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: For further details, see Notes for Table A3. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A6 First stage: Presence of a Treatment Table for Subgroups of the Population, Main Analysis Sample.

	(1) Female		(2) Low BW		(3) Firstborn		(4) Mother $\leq$ 23		(5) Hosp. birth		(6) Miss. father		(7) Mother $\geq$ mand. edu.		(8) Breastfed	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Born 1-3	0.563***	0.573***	0.567***	0.560***	0.610***	0.534***	0.519***	0.599***	0.593***	0.475***	0.572***	0.495***	0.583***	0.555***	0.593***	0.615***
	(0.008)	(0.008)	(0.006)	(0.027)	(0.008)	(0.007)	(0.009)	(0.007)	(0.006)	(0.012)	(0.005)	(0.023)	(0.007)	(0.008)	(0.009)	(0.008)
Ratio to full pop.	0.99	1.01	1.00	0.99	1.07	0.94	0.91	1.05	1.04	0.84	1.01	0.87	1.03	0.98	1.04	1.08
No. of obs.	44,646	42,690	76,631	4,017	38,868	47,886	33,590	53,164	68,237	19,099	82,591	4,745	44,342	39,061	30,820	41,324

Notes: Each cell shows estimates from separate regressions of the first stage estimated across subgroups of the population (denoted in the column headings). Regressions include fixed effects but exclude control variables. Breastfed equals one if the child was registered by a nurse as being breasted at one month of age. The ratio to full population refers to the value of the estimate in the subgroup compared to the value of the estimate in the full population. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A7 Robustness: ITT Effects on the SES and Health Index, Sample of First children of the Given Family in the Nurse Records.

	(1)	(2)	(3)	(4)	(5)
Panel A: SES Index					
Born 1-3	0.019	0.016	0.015	0.015	0.022
	(0.013)	(0.013)	(0.013)	(0.014)	(0.018)
No. of obs.	61,595	61,595	61,595	56,890	12,769
Panel B: Health Index					
Born 1-3	0.036***	0.033**	0.033**	0.032**	0.028
	(0.013)	(0.013)	(0.013)	(0.014)	(0.017)
No. of obs.	64,603	64,603	64,603	59,499	13,338
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See Table 6 for further details. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A8 Placebo: The Effect of Day of the Month of Birth on the SES and Health Indices in the Danish Towns Sample.

	(1)	(2)	(3)	(4)	(5)
Panel A: SES Index					
Born 1-3	0.011	0.011	0.011	0.006	0.006
	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)
No. of obs.	64,197	64,197	64,197	63,147	14,591
Panel B: Health Index					
Born 1-3	0.006	0.006	0.006	0.003	0.015
	(0.013)	(0.013)	(0.013)	(0.013)	(0.017)
No. of obs.	66,727	66,727	66,727	64,846	14,950
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables, with the exception of birth weight, as we do not observe nurse registrations in the Aarhus, Odense and Aalborg sample. We add a town fixed effect to our list of controls. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A9 Power Calculations for Firstgeneration Long-Run Outcomes.

Variable	Req. effect size	No. of obs.
Yrs. of educ.	0.082	85,330
Above compulsory educ.	0.014	85,330
Professionbachelor education	0.013	85,330
Higher education	0.015	85,330
University education	0.010	85,330
Earnings at 25	4291.093	85,741
Avg. Earnings 30-50	5859.463	85,035
Avg. Empl. age 30-50	0.009	84,100
Out of Labor Force, Age 50	0.013	78,595
Out of Labor Force, Age 55	0.019	41,828
Diabetes	0.007	87,336
Cardiovascular disease	0.014	87,336
Heart disease	0.007	87,336
Asthma	0.007	87,336
Cancer	0.008	87,336
Any mental health contact	0.011	87,336
Infection	0.013	87,336
Hosp. nights 20-29	0.627	87,336
Hosp. nights 30-39	0.624	87,336
Hosp. nights 40-49	0.609	87,336
Hosp. nights 50-59	0.510	87,336
Ever hospitalized (30-59)	0.013	87,336

Notes: The table shows the results of a simple power calculation, estimating the required ITT effect size for a statistically significant difference across groups (defined by the day of the month of birth) at a five percent level with 80 percent power. Income is measured in 2015 inflation adjusted DKK. The diagnoses record whether an individual has ever been diagnosed with the illness.

Table A10 ITT Effects on Education, Income and Labor Market Status.

	(1)	(2)	(3)	(4)	(5)
Panel A: Years of Education					
Born 1-3	-0.012	-0.016	-0.016	-0.021	-0.014
	(0.029)	(0.029)	(0.029)	(0.031)	(0.039)
No. of obs.	85,330	85,330	85,330	78,343	17,428
Panel B: Above Compulsory Education					
Born 1-3	0.002	0.001	0.001	0.000	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
No. of obs.	85,330	85,330	85,330	78,343	17,428
Panel C: Higher Education					
Born 1-3	-0.001	-0.002	-0.002	-0.002	-0.002
	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
No. of obs.	85,330	85,330	85,330	78,343	17,428
Panel D: University Education					
Born 1-3	-0.002	-0.002	-0.002	-0.003	-0.007
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
No. of obs.	85,330	85,330	85,330	78,343	17,428
Panel E: Earnings at 25 (2015 DKK)					
Born 1-3	-261.621	-132.550	-183.412	944.203	2145.744
	(1532.828)	(1527.284)	(1527.429)	(1629.848)	(2015.663)
No. of obs.	85,741	85,741	85,741	78,707	17,498
Panel F: Average Earnings 30-50 (2015 DKK)					
Born 1-3	1959.908	1829.872	1781.936	1994.575	2875.040
	(2078.669)	(2076.627)	(2076.822)	(2201.250)	(2730.072)
No. of obs.	85,035	85,035	85,035	78,054	17,352
Panel G: Avg Share in Employment Ages 30-50					
Born 1-3	0.006*	0.006*	0.006*	0.006*	0.006
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
No. of obs.	84,100	84,100	84,100	77,232	17,196
Panel H: Out of the Labor Force, Age 50					
Born 1-3	-0.007	-0.007	-0.007	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
No. of obs.	78,595	78,595	78,595	72,224	16,100
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 for further details. Robust standard errors in parentheses, * p<0.1 ** p<0.05 *** p<0.01.

Table A11 ITT Effects on Health Outcomes (Ever Diagnosed).

	(1)	(2)	(3)	(4)	(5)
Panel A: Diabetes					
Born 1-3	-0.003	-0.003	-0.003	-0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel B: Cardiovascular Disease					
Born 1-3	-0.005	-0.004	-0.004	-0.003	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel C: Heart Disease					
Born 1-3	-0.002	-0.001	-0.001	-0.002	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel D: Asthma					
Born 1-3	-0.005**	-0.005**	-0.005**	-0.007***	-0.008**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel E: Cancer					
Born 1-3	-0.005*	-0.005	-0.005	-0.006*	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel F: Mental Health Issues					
Born 1-3	-0.001	-0.001	-0.001	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel G: Infection					
Born 1-3	-0.001	-0.001	-0.001	-0.003	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables. Robust standard errors in parentheses.* p<0.1 ** p<0.05 *** p<0.01.

Table A12 ITT Effects on Health Outcomes (Hospitalizations).

	(1)	(2)	(3)	(4)	(5)
Panel A: Hospital Nights, Age 30-39					
Born 1-3	-0.229	-0.220	-0.216	-0.189	0.041
	(0.203)	(0.203)	(0.203)	(0.230)	(0.281)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel B: Hospital Nights, Age 40-49					
Born 1-3	-0.335	-0.332	-0.329	-0.487**	-0.313
	(0.222)	(0.222)	(0.222)	(0.247)	(0.297)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel C: Hospital Nights, Age 50-59					
Born 1-3	-0.128	-0.092	-0.088	-0.085	-0.317
	(0.190)	(0.189)	(0.189)	(0.217)	(0.269)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Panel D: Ever Hospitalized, Age 30-59					
Born 1-3	-0.007	-0.006	-0.006	-0.009*	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables. Robust standard errors in parentheses.* p<0.1 ** p<0.05 *** p<0.01.

Table A13 Robustness: ITT Effects on the Health Index, Excluding Asthma Diagnoses and Hospital Admissions.

	(1)	(2)	(3)	(4)	(5)
Panel B: Health Index					
Born 1-3	0.024**	0.022**	0.022**	0.021*	0.018
	(0.011)	(0.011)	(0.011)	(0.012)	(0.015)
No. of obs.	87,336	87,336	87,336	79,806	17,770
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See Table 5 for further details. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A14 ITT Effects on Education, Income and Labor Market Status, Sample of Females.

	(1)	(2)	(3)	(4)	(5)
Panel A: Years of Education					
Born 1-3	0.020	0.014	0.014	0.018	0.007
	(0.040)	(0.040)	(0.040)	(0.043)	(0.053)
No. of obs.	41,954	41,954	41,954	38,522	8,616
Panel B: Above Compulsory Education					
Born 1-3	0.011*	0.010	0.010	0.010	0.018*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)
No. of obs.	41,954	41,954	41,954	38,522	8,616
Panel C: Higher Education					
Born 1-3	-0.001	-0.002	-0.001	0.001	-0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)
No. of obs.	41,954	41,954	41,954	38,522	8,616
Panel D: University Education					
Born 1-3	-0.003	-0.003	-0.003	-0.004	-0.008
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
No. of obs.	41,954	41,954	41,954	38,522	8,616
Panel E: Earnings at 25 (2015 DKK)					
Born 1-3	-553.091	-406.517	-436.706	633.993	488.056
	(1947.452)	(1942.316)	(1942.424)	(2105.828)	(2616.144)
No. of obs.	41,987	41,987	41,987	38,543	8,600
Panel F: Average Earnings 30-50 (2015 DKK)					
Born 1-3	3898.510	3798.049	3745.357	2988.983	1584.724
	(2395.157)	(2390.050)	(2390.151)	(2564.418)	(3227.889)
No. of obs.	41,674	41,674	41,674	38,251	8,525
Panel G: Avg Share in Employment Ages 30-50					
Born 1-3	0.015***	0.015***	0.015***	0.014***	0.018***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
No. of obs.	41,218	41,218	41,218	37,855	8,448
Panel H: Out of the Labor Force, Age 50					
Born 1-3	-0.019***	-0.019***	-0.019***	-0.015**	-0.019**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)
No. of obs.	39,206	39,206	39,206	36,026	8,043
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 for further details. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A15 ITT Effects on Education, Income and Labor Market Status, Sample of Males.

	(1)	(2)	(3)	(4)	(5)
Panel A: Years of Education					
Born 1-3	-0.047	-0.049	-0.049	-0.065	-0.038
	(0.043)	(0.043)	(0.043)	(0.045)	(0.056)
No. of obs.	43,376	43,376	43,376	39,821	8,812
Panel B: Above Compulsory Education					
Born 1-3	-0.009	-0.009	-0.009	-0.010	-0.003
	(0.007)	(0.007)	(0.007)	(0.008)	(0.010)
No. of obs.	43,376	43,376	43,376	39,821	8,812
Panel C: Higher Education					
Born 1-3	-0.004	-0.004	-0.004	-0.005	-0.004
	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)
No. of obs.	43,376	43,376	43,376	39,821	8,812
Panel D: University Education					
Born 1-3	-0.001	-0.001	-0.001	-0.003	-0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
No. of obs.	43,376	43,376	43,376	39,821	8,812
Panel E: Earnings at 25 (2015 DKK)					
Born 1-3	653.647	731.187	651.908	1281.107	3686.190
	(2295.760)	(2285.279)	(2285.155)	(2482.721)	(3056.195)
No. of obs.	43,754	43,754	43,754	40,164	8,898
Panel F: Average Earnings 30-50 (2015 DKK)					
Born 1-3	857.793	681.822	623.952	891.458	3926.337
	(3303.054)	(3301.996)	(3302.802)	(3565.575)	(4371.480)
No. of obs.	43,361	43,361	43,361	39,803	8,827
Panel G: Avg Share in Employment Ages 30-50					
Born 1-3	-0.002	-0.002	-0.002	-0.002	-0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
No. of obs.	42,882	42,882	42,882	39,377	8,748
Panel H: Out of the Labor Force, Age 50					
Born 1-3	0.005	0.005	0.005	0.003	0.006
	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)
No. of obs.	39,389	39,389	39,389	36,198	8,057
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 for further details. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A16 ITT Effects on Education and Labor Market Index, Sample of Females and Males.

	(1)	(2)	(3)	(4)	(5)
Panel A: Education Index, Females					
Born 1-3	0.007	0.005	0.005	0.008	0.010
	(0.016)	(0.016)	(0.016)	(0.018)	(0.022)
No. of obs.	41,954	41,954	41,954	38,522	8,616
Panel B: Education Index, Males					
Born 1-3	-0.013	-0.014	-0.013	-0.020	-0.012
	(0.016)	(0.016)	(0.016)	(0.017)	(0.021)
No. of obs.	43,376	43,376	43,376	39,821	8,812
Panel C: Labor Market Index, Females					
Born 1-3	0.039***	0.040***	0.040***	0.033**	0.037*
	(0.015)	(0.015)	(0.015)	(0.016)	(0.021)
No. of obs.	39,107	39,107	39,107	35,941	8,015
Panel D: Labor Market Index, Males					
Born 1-3	0.001	0.001	0.000	0.001	0.010
	(0.018)	(0.018)	(0.018)	(0.019)	(0.023)
No. of obs.	39,317	39,317	39,317	36,133	8,044
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 for further details. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A17 ITT Effects on Health Outcomes (Ever Diagnosed), Sample of Females.

	(1)	(2)	(3)	(4)	(5)
Panel A: Diabetes					
Born 1-3	-0.006**	-0.006**	-0.006**	-0.003	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel B: Cardiovascular Disease					
Born 1-3	-0.009	-0.008	-0.008	-0.007	-0.010
	(0.007)	(0.007)	(0.007)	(0.008)	(0.010)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel C: Heart Disease					
Born 1-3	-0.004	-0.004	-0.004	-0.005	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel D: Asthma					
Born 1-3	-0.005	-0.005	-0.004	-0.007*	-0.009*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel E: Cancer					
Born 1-3	-0.007*	-0.007	-0.007	-0.008*	-0.005
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel F: Mental Health Issues					
Born 1-3	0.003	0.003	0.003	0.007	0.008
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Panel G: Infection					
Born 1-3	0.000	0.000	0.000	-0.004	-0.004
	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)
No. of obs.	42,690	42,690	42,690	39,021	8,715
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables. Robust standard errors in parentheses.* p<0.1 ** p<0.05 *** p<0.01.

Table A18 ITT Effects on Health Outcomes (Ever Diagnosed), Sample of Males.

	(1)	(2)	(3)	(4)	(5)
Panel A: Diabetes					
Born 1-3	-0.000	0.000	0.000	0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel B: Cardiovascular Disease					
Born 1-3	-0.002	-0.001	-0.001	0.001	0.001
	(0.007)	(0.007)	(0.007)	(0.008)	(0.010)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel C: Heart Disease					
Born 1-3	0.001	0.001	0.001	0.001	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel D: Asthma					
Born 1-3	-0.006*	-0.006*	-0.006*	-0.007**	-0.006
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel E: Cancer					
Born 1-3	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel F: Mental Health Issues					
Born 1-3	-0.006	-0.006	-0.006	-0.003	-0.005
	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Panel G: Infection					
Born 1-3	-0.004	-0.004	-0.003	-0.002	0.005
	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)
No. of obs.	44,646	44,646	44,646	40,785	9,055
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables. Robust standard errors in parentheses.* p<0.1 ** p<0.05 *** p<0.01.

Table A19 Robustness: ITT Effects on the SES and Health Indices, Sample of Survivors/Danish Residents in 2017.

	(1)	(2)	(3)	(4)	(5)
Panel A: SES Index					
Born 1-3	0.005	0.005	0.004	0.001	0.010
	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)
No. of obs.	76,468	76,468	76,468	70,297	15,636
Panel B: Health Index					
Born 1-3	0.030***	0.029***	0.029***	0.031***	0.031**
	(0.011)	(0.011)	(0.011)	(0.012)	(0.015)
No. of obs.	77,299	77,299	77,299	71,056	15,814
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See Table 5 for further details. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A20 Placebo: ITT Effects on Probability of Emigration and Death in the Town Sample (Outside Copenhagen), 1977-2020.

	Ever Emigr. (1)	Death (2)
Born 1-3	-0.001	0.002
	(0.005)	(0.003)
MDV	0.147	0.067
No. of obs.	66,727	66,727

Notes: Regressions are based on a specification including year of birth and day of the week of birth fixed effects. See Notes for Table A8 for further details. Individuals enter the sample if observed in the administrative data at Statistics Denmark at least once between 1977-2017. Deaths are measured in the death registry for the 1970-2020 period. However, as individuals in our town sample need to survive to 1977 to be observed in the register data at Statistics Denmark, for them we only can observe deaths in the 1977-2020 period. This period deviates from the period observed for our nurse record sample, where we observe individuals' unique personal identifier prior to 1977 and thus can merge on deaths for the 1970-2020 period (as we have obtained the personal identifier directly from the CPR registry covering the 1968 period and after (see section 3)). Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A21 Mechanisms: ITT Effects on Childhood Height.

	- - Height at - -				
	Age 3	Age 6	Age 7	Age 10	Age 13
Born 1-3	0.829**	1.099*	0.442	0.098	-0.453
	(0.381)	(0.578)	(0.407)	(0.434)	(0.544)
MDV	95.81	118.43	122.86	138.13	155.48
No. of obs.	1,206	1,090	1,773	2,400	2,348
Incl. Low Birthweight Control					
Born 1-3	0.715*	1.013*	0.372	0.038	-0.489
	(0.384)	(0.579)	(0.411)	(0.433)	(0.540)
MDV	95.81	118.43	122.86	138.13	155.48
No. of obs.	1,206	1,090	1,773	2,400	2,348

Notes: Each column presents estimates from a separate regression of the respective outcome on an indicator for the child being born in the first three days of the month for the samples of CPC children observed in the nurse records. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A22 Mechanisms: ITT Effect on Indicators for full Uptake of Recommended Vaccinations and Indicators for Ever Having Used Antibiotics (Mother-Reported).

	Vacc., Age 3	Vacc., Age 6	Antib., Age 1-3	Antib., Age 3-6
Born 1-3	0.048	0.022	-0.093***	-0.145***
	(0.034)	(0.032)	(0.036)	(0.037)
MDV	0.60	0.44	0.48	0.48
No. of obs.	2,402	2,716	2,273	1,888
Incl. Low Birthweight Control				
Born 1-3	0.047	0.022	-0.089**	-0.145***
	(0.034)	(0.032)	(0.036)	(0.037)
MDV	0.60	0.44	0.48	0.48
No. of obs.	2,402	2,716	2,273	1,888

Notes: see Notes for Table A21. * p<0.1 ** p<0.05 *** p<0.01.

Table A23 Mechanisms: ITT Effects on Child Ever Having Been Admitted to Hospital and of Ever Having Been Diagnosed with Infectious Disease (Mother-Reported).

	Hosp, 1-3	Hosp, 3-6	Inf. disease, 1-3	Inf. disease, 3-6
Born 1-3	-0.020	-0.035	-0.045	-0.037
	(0.030)	(0.030)	(0.029)	(0.029)
MDV	0.27	0.33	0.30	0.31
No. of obs.	2,402	2,716	2,704	2,743
Incl. Low Birthweight Control				
Born 1-3	-0.015	-0.035	-0.044	-0.037
	(0.030)	(0.030)	(0.029)	(0.029)
MDV	0.27	0.33	0.30	0.31
No. of obs.	2,402	2,716	2,704	2,743

Notes: see Notes for Table A21. * p<0.1 ** p<0.05 *** p<0.01.

Table A24 Mechanisms: ITT Effect on Age at Completion of Developmental Milestones (Mother-Reported), Z-Scores.

	Language	Motor	Eating	Dressing	Soc. Inter.	Toilet	Comb. score
Born 1-3	-0.053	-0.105	-0.153**	0.024	-0.175**	-0.019	-0.121*
	(0.076)	(0.072)	(0.076)	(0.081)	(0.080)	(0.075)	(0.071)
MDV	0.01	0.01	0.01	-0.00	0.02	0.00	0.01
No. of obs.	2,001	2,180	1,946	1,698	2,021	2,134	2,256
Incl. Low Birthweight Control							
Born 1-3	-0.048	-0.093	-0.149**	0.027	-0.166**	-0.010	-0.110
	(0.076)	(0.072)	(0.076)	(0.081)	(0.080)	(0.074)	(0.071)
MDV	0.01	0.01	0.01	-0.00	0.02	0.00	0.01
No. of obs.	2,001	2,180	1,946	1,698	2,021	2,134	2,256

Notes: see Notes for Table A21. * p<0.1 ** p<0.05 *** p<0.01.

Table A25 ITT Effects on Focal Children's Mothers: Fertility and Labor Market Outcomes.

	(1)	(2)	(3)	(4)	(5)
Panel A: Age at Birth					
Born 1-3	0.044	0.071	0.072	0.085	-0.001
	(0.079)	(0.078)	(0.078)	(0.086)	(0.106)
No. of obs.	47,882	47,882	47,882	44,004	9,828
Panel B: Spacing (Days)					
Born 1-3	3.484	0.474	1.126	-5.984	-9.282
	(16.519)	(16.456)	(16.464)	(17.763)	(21.885)
No. of obs.	36,561	36,561	36,561	33,542	7,438
Panel C: No. of Children					
Born 1-3	-0.012	-0.011	-0.011	-0.017	-0.001
	(0.014)	(0.014)	(0.014)	(0.015)	(0.019)
No. of obs.	47,882	47,882	47,882	44,004	9,828
Panel D: Average Earnings 40-65 (2015 DKK)					
Born 1-3	-186.876	-1022.071	-1062.895	-246.853	-1158.010
	(1841.605)	(1819.284)	(1820.169)	(1972.158)	(2462.026)
No. of obs.	47,058	47,058	47,058	43,240	9,652
Panel E: Average Employment 40-65					
Born 1-3	-0.003	-0.005	-0.005	-0.003	-0.008
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
No. of obs.	46,505	46,505	46,505	42,777	9,550
Cohort FE	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	Yes	Yes	Yes
Pre-treatment controls	No	No	No	Yes	Yes
Only born 1-7	No	No	No	No	Yes

Notes: See *Notes* for Table 5 on control variables. Standard errors are robust. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A26 ITT Effects on Fertility of Focal Children.

	Childless	No. of Children	Age at First Child
Born 1-3	-0.001	-0.015	-0.042
	(0.005)	(0.012)	(0.071)
MDV	0.22	2.12	27.81
No. of obs.	87,336	68,522	68,503

Notes: Each cell presents the estimate of a separate regression. All regressions include fixed effects for year of birth and day of the week of birth, as well as child sex (all defined for the focal child). Sample includes all focal individuals (m/f) with non-missing data on initial health status control variables. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A27 ITT Effects on Fertility of Focal Children, by Sex of Focal Child.

Outcome	Childless		No. of Children		Age at First Child	
	M	F	M	F	M	F
Born 1-3	0.006	-0.007	-0.040**	0.007	-0.018	-0.015
	(0.007)	(0.006)	(0.017)	(0.016)	(0.104)	(0.093)
MDV	0.26	0.17	2.12	2.11	29.20	26.51
No. of obs.	44,646	42,690	33,027	35,495	33,021	35,482

Notes: Samples includes all focal children (M: male; F: female) with non-missing information on initial health status. For more details see Notes for Appendix Table A26. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A28 ITT Effects on Fertility of Focal Children, by Low Health Status.

Outcome	Childless		No. of Children		Age at First Child	
Low BW Status	0	1	0	1	0	1
Born 1-3	0.011*	-0.013	-0.039**	0.005	0.027	-0.090
	(0.007)	(0.008)	(0.016)	(0.021)	(0.102)	(0.121)
MDV	0.21	0.23	2.10	2.13	28.26	27.16
No. of obs.	47,357	32,753	37,502	25,411	37,493	25,401

Notes: For more details see Notes for Appendix Table A26. Sample includes all focal individuals with non-missing information on initial health status. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A29 The Effects of Extended Nurse Care on Health at Birth, Second Generation Outcomes (ITT).

	(1) Birth Weight	(2) Low BW	(3) Preterm
A. All Children of Focal Mothers			
Born 1-3	-2.514	-0.001	0.003
	(7.545)	(0.003)	(0.003)
MDV	3407.34	0.06	0.06
No. of obs.	71,145	71,145	70,016
B. All Children of Focal Fathers			
Born 1-3	1.232	-0.002	0.002
	(7.761)	(0.003)	(0.003)
MDV	3436.54	0.06	0.06
No. of obs.	66,699	66,699	65,966

Notes: Each cell presents the estimate from a separate regression. The samples includes all second generation children of focal mothers or fathers, respectively. Controls include focal children's yob, day of the week of birth fixed effects. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table A30 The effect of Extended Nurse Care on Birth Outcomes in the Second Generation. Heterogeneity by Focal Child Initial Health (Low Health Status) (ITT)

Second Gen.	Birth Weight (g)		Low BW		Preterm	
First Gen. Low Health	0	1	0	1	0	1
A. All Children of Focal Mothers						
Born 1-3	-13.736	4.068	0.001	-0.006	0.007	-0.002
	(10.684)	(12.659)	(0.004)	(0.005)	(0.004)	(0.005)
MDV	3439.59	3361.05	0.05	0.07	0.06	0.07
No. of obs.	38,071	27,185	38,071	27,185	37,523	26,712
B. All Children of Focal Fathers						
Born 1-3	-5.385	17.419	0.000	-0.007	0.004	-0.001
	(10.809)	(13.454)	(0.004)	(0.005)	(0.004)	(0.006)
MDV	3458.93	3403.96	0.05	0.06	0.06	0.07
No. of obs.	36,637	24,596	36,637	24,596	36,245	24,305

Notes: Each cell shows estimates from separate regressions (for second generation birth outcomes) across subgroups of the population (defined by the low health status of the focal individual, who is now a parent). The samples consists of all children born to either focal mothers or fathers, respectively. The estimates come from our heterogeneity specification including fixed effects for focal children's year of birth, day of the week of birth. MDV is the mean of the dependent variable in the control group for the relevant subgroup. * p<0.1 ** p<0.05 *** p<0.01.

Table A31 The Effect of Extended Nurse Care on Education and Earnings in the Second Generation (Children born prior to 1991). Heterogeneity by Focal Child Initial Health

Second Gen.	Yrs. of Edu.		Above Comp. Edu.		Avg. Earn., age 20-29	
First Gen. Low Health	0	1	0	1	0	1
Born 1-3	-0.182***	0.025	-0.020**	0.015	1809.693	1722.406
	(0.060)	(0.065)	(0.010)	(0.011)	(2370.607)	(2633.226)
MDV	14.30	13.72	0.80	0.73	164262.91	158651.49
No. of obs.	21,447	17,646	21,447	17,646	21,599	17,831

Notes: Each cell shows estimates from separate regressions (for second generation birth outcomes) across subgroups of the population (defined by the initial health status of the focal individual, who is now a parent). The samples consists of all children born to focal individuals prior to or in 1990. If a child has both parents (focal children) in our data, we keep the mother spell to assign treatment and low health status to the second generation child. The estimates come from our heterogeneity specification including fixed effects for focal children's year of birth and day of the week of birth, as well as focal child sex. MDV is the mean of the dependent variable in the control group for the relevant subgroup. * p<0.1 ** p<0.05 *** p<0.01.

B The Coverage of the Copenhagen Nurse Program and the Coverage of the Nurse Records

In this section, we describe (i) the overall coverage of the Copenhagen nurse program and (ii) the coverage of our linked nurse records among the Copenhagen children of the 1959-1967 cohorts, who were eligible for the program. The main complication for these analyses is that we do not observe the number of (relevant) resident children in Copenhagen at the time in neither aggregate or individual level data. While resident children are the relevant population, in our administrative data we only observe the number of Copenhagen births: Until 1978, the default for birth registrations in Denmark was the parish of birth, with parishes being nested in municipalities. As one exception, hospital births were registered with a hospital code. While home births remained the norm in many parts of the country until the early 1970s, in and around larger towns (especially the capital Copenhagen) hospital births were more common. Especially women with health and social disadvantages, often not residing in Copenhagen, gave birth at hospitals.

Using different data sources, we obtain the following main results: Aggregate yearly statistics from the time indicate a broad initial coverage of the nurse program among Copenhagen births (and likely residents). Up to one third of infants who initially enter the nurse program, discontinue the program during the first year. We find that the number of children that we observe in the nurse records corresponds nicely to aggregate figures of program initiation, with discepancies likely being due to measurement issues (stock vs flows of children in the nurse program). Furthermore, looking at our individual level matched records, those cover roughly two thirds of the children actually *born* in Copenhagen. Factoring in that hospital births may not all reflect Copenhagen residency and that inner-Danish migration was common (i.e., we have a considerable share of infants not born in Copenhagen but living in Copenhagen and thus being eligible for the first year nurse program), our analysis supports that the nurse records cover the vast majority of relevant infants and their families.

Aggregate records on coverage of the Copenhagen nurse program Appendix Table B1 shows aggregate statistics collected by Copenhagen officials at the time (Copenhagen Statistical Office, various years) and figures from our matched nurse data: The number of Copenhagen births, the number of infants ever entering the nurse program (resident children), the number of infants discontinuing the nurse program before their first birthday, the number of infants that we observe with a matched nurse record, and the share of infants in our nurse data among all supervised infants (in the aggregate data). Data comes from different sources (e.g. aggregate tables on births and children entering the nurse program).

There are two main findings from the table: First, the first two columns suggest that the Copenhagen program initially reached out to a vast majority of infants thus being a universal program.

Second, the infants that we study correspond nicely to the aggregate records on infants entering treatment in the years under consideration. This confirms a very good coverage of the nurse records among ever treated individuals. Over the years, coverage rates of our records that exceed 100 percent are observed. This discrepancy must be attributed to different data sources in the aggregate data, potentially changing definitions of CPH births (including or excluding hospital births by non-CPH residents) over time, and discrepancies related to "first in the year stock" vs flow measures.

A final observation relates to the aggregate figures on discontinuation of supervision in the first year program. In the light of up to 30 percent of infants leaving the first year program (mainly due to mobility), our first stage results appear reasonable.

Table B1 Copenhagen Births and Nurse Record Coverage, Copenhagen Statistical Yearbooks.

Year	CPH Births	Infants Entering Supervision	Supervision Discontinued	Nurse Records	Nurse Records Among Supervised Infants (%)
1959	9,316	8,690	1,460	8,620	99.2
1960	9,876	8,641	1,891	9,437	109.2
1961	9,800	8,951	2,178	10,120	113.1
1962	9,889	8,902	2,294	10,657	119.7
1963	10,568	10,708	3,059	11,411	106.6
1964	10,430	10,548	3,166	11,322	107.3
1965	10,430	10,566	3,130	10,749	101.7
1966	10,593	10,697	3,357	9,932	92.8
1967	9,690	9,771	3,166	10,031	102.7

Notes: The table shows aggregate statistics from various sources and descriptive statistics from the nurse records. The number of Copenhagen births and the number of children entering and leaving nurse supervision comes from two sets of statistical accounts in the Copenhagen Statistical Yearbooks (Copenhagen Statistical Office, various years). The number of nurse records is the number of relevant records in our external (individual level) data.

Individual level data on place of birth and the coverage of the scanned nurse records Appendix Table B2 shows data on the types of birth registration for the 1959-1967 Copenhagen births in the administrative data: It displays the type of birth registrations for all Copenhagen births (100,854), the Copenhagen births unmatched and matched, respectively, to our nurse records. Of the individuals born in Copenhagen, we can match 71,596 (71 percent of the 100,854 births) to a nurse record.[48] However, non-Copenhagen residents are likely an important part of the explanation for us not finding all Copenhagen births in the nurse data: Among the infants born in Copenhagen but not covered in the nurse data, 46 percent are born in a hospital in the city. Thus it is likely that those children did not actually reside in Copenhagen, but that their mothers came to the capital to give birth (either due to medical complications that made a hospital birth necessary or to give birth and potentially give up their child for adoption). Similarly, starting out with our nurse data, we find that 82 percent of those infants were born in Copenhagen. The remainder of infants in the nurse data were mainly born in adjacent municipalities, among them Frederiksberg, Tårnby,

[48]Factoring in the unmatched records (those without a personal identifier assigned), we can create an upper bound for the coverage of the records among Copenhagen births: Assuming that all the unmatched records are relevant Copenhagen births, we find that the data cover up to 71,596 + 4,094 = 75,690 (75 percent) of all Copenhagen births.

Table B2 Share (%) of Copenhagen Births (Administrative Data) covered in the Nurse Records.

	All CPH Births	Unmatched CPH Births	Nurse Record CPH Births
Parish	67.33	53.49	72.98
Hospital	32.29	45.84	26.75
Municipality	0.38	0.66	0.26
Total	100.00	100.00	100.00
N	100854	29258	71596

Notes: The table shows the type of birth registration for all Copenhagen births, for unmatched Copenhagen births, and for matched individuals covered in the nurse records. The birth registration code comes from the administrative data and thus only covers individuals who have survived to 1977.

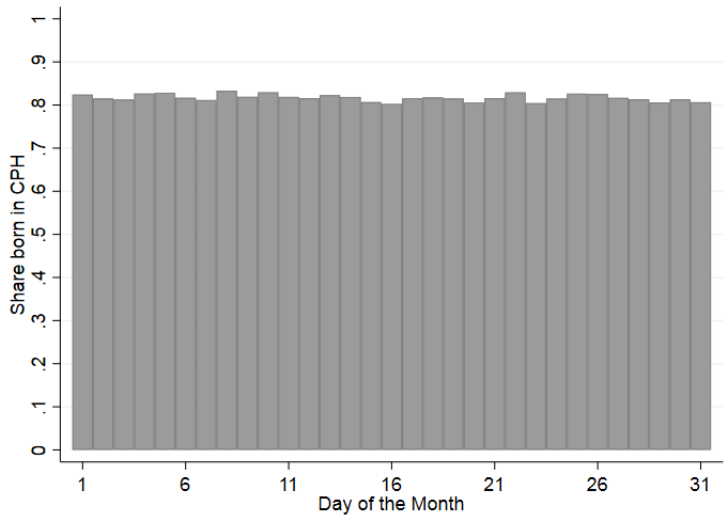


Fig. B1 Share of Copenhagen Births in the Matched Nurse Data Across Days of the Month, 1959-1967.

and Gentofte (those municipalities account for 57, 11, and nine percent, respectively, of the children identified in the nurse records but not born in Copenhagen). Thus the individual level data on birth registrations supports our claim that the nurse records have a very good coverage among the infants entering the nurse program.

Finally, Appendix Figure B1 verifies that the share of Copenhagen births in our sample is stable across days of the month (suggesting no patterns of hospital births that vary across days of the month and could interfere with our design that exploits this dimension).

C Additional Details on Data

ICD Codes used in the Health Measures

We use data from the Danish Inpatient Register to measure discharge diagnoses in the 1977-2018 period. In 1995, the diagnoses scheme applied changed from the ICD 8 scheme to use the ICD 10 scheme. As we cannot easily distinguish the types of diabetes in the ICD 8 scheme, we only use the ICD 10 data for diabetes (given that diabetes is an absorbing state and that prevalence and diagnoses increase with age, we think this decision is reasonable). For mental health condiitions we code the indicator as one for an individual who is ever observed in the mental health discharge data with any mental health diagnosis.

Diagnosis (Indicator)	Underlying ICD 8 and ICD 10 codes
Diabetes	DE11, DE13, DE14
Cardiovascular Disease	390-458; DI00-DI99
Heart Disease	410-414; DI20-DI25
Asthma	493; DJ45-DJ46
Cancer	140-209 (excl. 173); DC00-DC97 (excl. DC44)
Infections	000-136; DA00-DA99, DB00-DB99

Table C1 Hospital Diagnoses Measures and Underlying ICD Codes

D Details on Transcription Methods and Validity

This Appendix describes how we extract data from the scanned nurse records. We also assess the performance of the applied methods by validating them on the manually transcribed subsets of the nurse records.

D.1 Detection and Transcription of the Treatment Table

Our goal is to detect the presence of the treatment table among all pages of a given nurse record. We also transcribe the contents of the treatment tables, to assess the completeness of registrations. To detect the treatment table, we apply an ML layout classification procedure that is detailed in Dahl et al. (2021). For convenience, we provide a summary of the method and results below.

A typical nurse record with a treatment table is shown in Figure 1 where, in this case, the treatment table is present on the third page. The page number containing the table varies, and thus an approach to always consider the third page is not helpful. We use an unsupervised ML method to organize the pages according to their appearance and layout such that similar looking pages are grouped. This is especially applicable to the nurse records as the scan quality and alignment is very uniform across the collection and hence any visual deviations between the pages reflect their content rather than scan artifacts. One key advantage of the unsupervised method is that we do not need to collect any training data.

We start by constructing a lower-dimensional representation of the images by extracting a feature vector from each page in every nurse record. For feature extraction, we pursue a transfer learning approach and re-use a VGG-16 neural network (Simonyan and Zisserman, 2015) pre-trained to perform classification on ImageNet.[49] We strip the classification layers of the VGG-16 network and rely on the 512-dimensional vector representation that can be

[49]ImageNet (Deng et al., 2009) is a dataset that is often used for benchmarking and testing image classification models. It contains more than a million images across 1000 different categories. The pre-trained network parameters we use are available through the torchvision package, see `https://pytorch.org/vision/stable/index.html`

extracted from the final convolutional layer of the network. This representation serves as our feature vector for a given page and we extract one such vector for every image (page) in the collection of 261,926 pages. Next, all the feature vectors are clustered using the DBSCAN algorithm (Ester et al., 1996) which produces 37 clusters. We manually annotate the clusters by randomly sampling ten pages from each cluster, 370 pages total, and assigning a label to each of them. In this process, we use t-distributed Stochastic Neighbour Embedding (t-SNE) (van der Maaten and Hinton, 2008) to visualise the feature vectors and clusters. The t-SNE method produces embeddings of the feature vectors in two dimensional space that preserves their structure, i.e. points that are close in feature space also tend to be close in the low-dimensional embedding space. Figure C1 shows the results, where a clear separation of the different clusters, and thus also the different page types, is apparent.

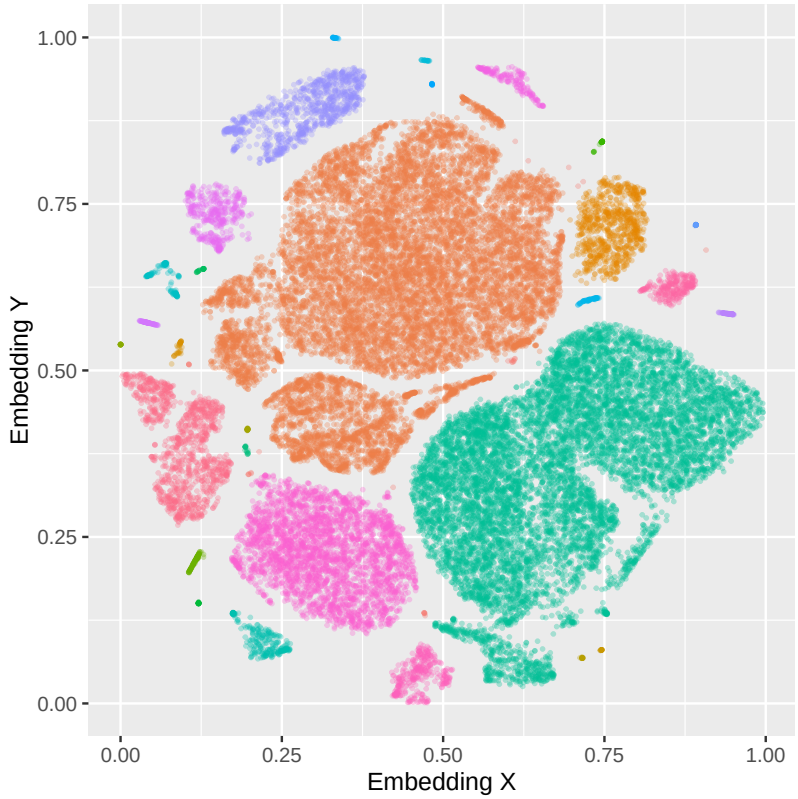


Fig. C1 2D t-SNE visualisation of the feature space of the record pages.

Notes: Each point represents a record page and the colours correspond to the labels assigned by the clustering algorithm. Pages with similar layout cluster together. The embeddings have been subsampled to reduce cluttering, so only 30,000 randomly sampled embeddings are displayed. There is a total of 37 clusters which are manually annotated. The treatment pages are contained in four clusters.

Table C1 Confusion matrix for the treatment detection model.

	ML detection	
Ground truth	Treated	Not Treated
Treated	234	0
Not Treated	0	3766

Notes: The table shows a confusion matrix for our treatment detection model. The frequencies are based on a randomly sampled and manually reviewed test set of 4000 records (10,914 pages). The treatment detection model does not rely on any segmentation, but detects the presence of the whole page containing the treatment table.

We evaluate the clustering procedure on a test dataset of 4000 randomly selected records. For each record, we review all pages and check whether they contain the treatment table. We then compare the predictions from the clustering approach to the manually assigned labels. The results of this validation is summarized in Table C1. We find no errors, i.e. all 4000 records are classified correctly by our clustering approach despite heavy class imbalance. In this light, we are confident in the quality of the clustering approach to identify which record pages – and hence which records – contain the treatment table.

We use a supervised convolutional neural network to classify the contents of the treatment table (specifically the content of each cell of these tables). We manually transcribed 2000 random cells and use these to train a supervised convolutional neural network to classify the contents of the cells by considering three types of content (manually written entry, empty, and machine written cross, indicating that the cell should not be filled out). We test the accuracy of this model on a held-out test set, where it achieves an accuracy of 99.4%.

D.2 Detection and Transcription of Other Fields

In addition to identifying which records contain a treatment table and transcribing their contents, we also transcribe the cells of tables detailing information up to one year of age (tables of the form on page one of Figure 1 associated with the universal portion of the nurse visiting program).

To transcribe the contents of the tables, we use a three-steps approach, similar to that

described in Dahl et al. (2021). First, we identify the page where the table in question is present. The tables of the form on page one of Figure 1 are simple to detect, as they are most often on the first page in each record.

Second, the fields of data, which we refer to as "cells", in a table, such as the birth weight, are cut out as small images. The second step is performed on the first page of each record trying to segment the tables. As the procedure relies on the tables being present, it fails in cases where this is not the case. For these cases, we consider subsequent pages of the records to find the tables.

In order to cut out images, we use point set registration. A thorough discussion of this approach (used in another, but similar setting) is provided in Dahl et al. (2021). Here, we provide a brief explanation of the approach. It relies on aligning points between an image and a template (Besl and McKay, 1992). The motivation for using point set registration stems from the tabular structure of the documents we digitize. This allows us to extract points corresponding to the intersections between the vertical and horizontal lines of the tables. To extract the lines, we use an approach similar to what is described in Szeliski (2010). We then identify the points corresponding to the intersections of the lines by means of Harris corner detection (Harris and Stephens, 1988). This approach allows us to align the points across each image with a set of template points, which we do using coherent point drift (Myronenko and Song, 2010). This yields a transformation matrix for each image, which we use to transform the image such that we can crop each "cell" of the table on the image as a small image.

Finally, we use neural networks to classify the contents of each cell, using an approach similar to that in Dahl et al. (2022). The neural networks need training data, which we obtain through manual transcription of a subset (eight percent) of the fields of the nurse records. In total across all neural networks, we use 538,821 labelled fields of data to train and tune the neural networks to transcribe millions of fields of data. As our segmentation is not perfect, and the neural networks are not able to transcribe wrongly cropped cells such

as when the data of interest is not present in the crop, we allow them to predict that the segmentation step has failed. To provide training data for this, we supplement our training data with such cases, which we determine not through manual labelling but by considering records where the determinant of the transformation matrix is abnormally small or large.[50] We use five percent of this data as validation data to monitor the performance and tune the learning rate and strength of regularization (weight decay and data augmentation). The training schedule (length and learning rate decay) is mostly similar to standard training on ImageNet (Deng et al., 2009), such as in Xie et al. (2017).

We use multiple neural networks depending on the cells they work on. The neural networks are similar, and they are trained similarly. All code is written in Python using PyTorch (Paszke et al., 2019). Each neural network uses a ResNet-50 with bottleneck building blocks (He et al., 2016) as its feature extractor; the weights of the PyTorch version of ResNet-50 pretrained on ImageNet (Deng et al., 2009) are used as the initial weights. The neural networks differ only insofar as their classification heads differ. Some of the neural networks perform standard classification and others perform classification of sequences. When sequences are used, a method similar to the one described in Goodfellow et al. (2013) is used, with the exception that the sequence length is never estimated. The weights of the heads are randomly initialized.

All neural networks are trained using stochastic gradient descent with momentum of 0.9, weight decay of 0.0005, and Nesterov acceleration based on the formula in Sutskever et al. (2013). The batch size used is 256 and the initial learning rate is $\frac{0.05 \texttt{ batch-size}}{256}$. The networks are trained for 100 epochs and the learning rate is divided by ten every 30 epochs. As loss function, we use the negative log likelihood loss for networks performing standard classification. When training networks on sequences, we treat the problem similar to a multi-task learning problem. The loss of each classification head is the negative log likelihood loss of the head, and the total loss is the mean of the negative log likelihood loss of each head.

[50] The determinant of the transformation matrix is expected to be close to one, to retain the general shape of the overall image. If it is significantly smaller or larger, a part of the segmentation process has failed.

Images are resized differently between the neural networks, in a way to retain the approximate size of the cells in the record from which we define our template points for segmentation. The images are normalized using the ImageNet means and standard deviations (to normalize similarly to the pretrained ResNet-50 feature extractor). During training, we use image augmentation in the form of RandAugment with $N = 3$ and $M = 6$ (Cubuk et al., 2020), based on the implementation in Kim (2020).

We train different neural networks as they transcribe very different data. We use different neural network to transcribe children's weights, dates, child lengths, cells with single-digit values, total duration of breastfeeding, first names (of nurses), and last names (of nurses). We describe the heads, data, and image resizing of each separately below, as these are the settings that differ between the networks.

Weight The neural network transcribing weights consists of a classifier with six heads, the first with three nodes and the remaining with 11 nodes. The first output layer indicates whether the cell is {`normal`, `empty`, `failed segmentation`}, and the remaining five output layers each indicate a digit in the weight (with the eleventh option indicating "empty"). This allows the neural network to produce predictions in the form of the cell being empty, the segmentation in step 2 having failed, and any integer (up to 99,999) representing a weight. Some weights, however, are censored from above, as two types of scales were used, the first of which had a limit of ten kg. and the second of which had a limit of 12 kg. This is a limitation of the original data source and has nothing to do with the transcription; however, it means that the weights of some of the children at age nine months and many of the children at age 12 months provide only a censored estimate of their weights. This neural network transcribes birth weight and weight at one, two, three, four, six, nine, and 12 months of age. We resize the images to a height of 63 pixels and a width of 302 pixels.

Date The neural network transcribing dates consists of a classifier with four output layers, the first with two nodes, the second with four nodes, the third with 11 nodes, and the

fourth with 13 nodes. The first output layer indicates whether the cell is {`normal`, `failed segmentation`}, the second output layer the first digit of the day (with the fourth option indicating "empty"), the third output layer the second digit of the day (with the eleventh option indicating "empty"), and the fourth output layer the month (with the thirteenth option indicating "empty"). This allows the neural network to produce predictions in the form of the cell being empty, the segmentation in step 2 having failed, and any date consisting of a day and a month (which also allows for some invalid dates, such as February 31). Although the year is sometimes available, in most cases only the day and month is present in a cell. As such, to use these predictions we post-process the predicted dates by adding the year, which we increment by one when we pass from an end-of-year month to a start-of-year month. This neural network transcribes exact date of visits at one, two, three, four, six, nine, and 12 months of age. We resize the images to a height of 63 pixels and a width of 212 pixels.

Length The neural network transcribing lengths consists of a classifier with two output layers, both with 11 nodes. The eleventh option for each layer indicates `failed segmentation`, and the remaining ten options indicate respective digits. This allows the neural network to produce predictions in the form of the segmentation in step 2 having failed and any integer (up to 99) representing a length. We resize the images to a height of 93 pixels and a width of 198 pixels.

Single-digit The neural network transcribing single-digit cells consists of a classifier with two output layers, the first with three nodes and the second with ten nodes. The first output layer indicates whether the cell is {`normal`, `empty`, `failed segmentation`} and the second output layer an integer between 0 and 9. This neural network transcribes all columns with single digit values of the visit at one, two, three, four, six, nine, and 12 months of age.[51] This allows the neural network to produce predictions in the form of the cell being empty, the segmentation in step 2 having failed, and any integer 0-9. Most of the cells are always

[51]This amounts to 112 cells in total for each record. Of these, we use 49 in this paper.

restricted to 1-2 or 1-3, but we train a joint neural network on all cells regardless to increase the size of the training data, and most importantly to alleviate a problem with the majority of the cells not having any labelled data.[52] We resize the images to a height of 64 pixels and a width of 125 pixels, and in this case use a higher batch size of 512.

Duration of breastfeeding The neural network transcribing duration of breastfeeding consists of a single output layer with 14 nodes. The fourteenth option indicates `failed segmentation`, and the remaining 13 nodes indicate 0-12 months (the duration). This allows the neural network to produce predictions in the form of the segmentation in step 2 having failed and any integer 0-12. We resize the images to a height of 93 pixels and a width of 293 pixels.

First name The neural network transcribing first name of nurses consists of 15 output layers, the first with two nodes and the remaining with 30 nodes. The first output layer indicates whether the cell is {`normal`, `failed segmentation`}, and the remaining 14 layers predicts letters a-å (29 letters) as well as an empty option. This allows the neural network to produce predictions in the form of the segmentation in step 2 having failed and any name up to 14 letters (this is the longest name in our training data), including an option of the cell being empty. We resize the images to a height of 95 pixels and a width of 680 pixels.

Last name The neural network transcribing last name of nurses consists of 15 output layers, the first with two nodes and the remaining with 30 nodes. The first output layer indicates whether the cell is {`normal`, `failed segmentation`}, and the remaining 14 layers predicts letters a-å (29 letters) as well as an empty option. This allows the neural network to produce predictions in the form of the segmentation in step 2 having failed and any name up to 14 letters (this is the longest name in our training data), including an option of the cell being empty. We resize the images to a height of 95 pixels and a width of 680 pixels.

[52]To assure our performance metrics are correct, we create a test set by manually labelling 11,200 cells, i.e. 100 of each type.

Performance of transcription We evaluate the performance of our transcription by assessing its performance on a held-out test set. Table C2 shows results across different fields. Note that most fields consists of multiple cells such as, e.g., the row showing the performance of weights consisting of weights measured at birth and one, two, three, four, six, nine, and 12 months of age. The accuracy is measured as the accuracy of the full sequence of predictions being correct, i.e. getting 13 out of 14 letters in a name correct provides no contribution to our measure of accuracy. As such, the character accuracy is significantly higher for most cells.[53]

We also show performance metrics when we exclude the cases most easily transcribed. Specifically, cells which are empty and cells where our segmentation step has failed are often significantly easier to transcribe, as the content of the cell in either case is very easy to detect (it being either empty or not containing the usual information). Further, for our analyses we are interested in using the data of cells which are neither empty nor with failed segmentation. As such, Table C2 also shows the performance of our method for the subset of non-empty cells, the subset of cells with successful segmentation, and the subset of cells which are non-empty and successfully segmented. The share of empty cells vary significantly between the different fields, while the segmentation process, being performed on a full image rather than the small images of cells we crop, is stable across different fields of data. For birth weight, we obtain measurements for 79,348 (92 percent) of our matched sample (86,553), while for other fields of data (such as later weight measurements) the share of cells with usable data is significantly lower (see Table 3).

[53]We do not believe that character accuracy is a particularly interesting metric, as we need the full sequence transcribed accurately: A weight of 6000 grams is significantly different than a weight of 5000 grams, and attributing such an error as being 75 percent correct is in our view not the correct metric for our task, even if three out of four digits are correctly transcribed. As such, our metric is more difficult than the commonly reported character accuracy.

Table C2 Automated transcription performance.

	Transcription accuracy (%)			
	All	Non-empty	Successful crop	Successful non-empty crop
Care and cleanliness	97.8	97.9	97.4	97.5
Date	96.9	96.6	96.4	95.9
Duration breastfeeding	98.7	98.7	98.0	98.0
Home economic status	97.6	97.7	97.4	97.4
Home harmony	97.8	97.5	97.7	96.9
Length	98.5	98.5	98.2	98.2
Mother mental capacity	98.1	98.3	97.9	98.0
Mother physical capacity	98.3	98.4	98.3	98.6
Nursery or kindergarten	97.6	97.8	97.2	97.4
Nutrition	97.2	97.4	96.7	96.8
Weight	96.8	96.7	96.1	96.0
Nurse name (first)	98.3	98.0	96.6	92.2
Nurse name (last)	99.1	99.1	98.5	97.0

Notes: The table shows the accuracy (in percent) of the ML transcriptions for separate types of fields, measured on an independent test set not part of the data used to train our neural networks. The second column shows the accuracy of the full test sample, and the subsequent columns of more difficult subsamples. The third column shows the accuracy when excluding empty fields, which are often easier to transcribe. The fourth column shows the accuracy when excluding fields where our segmentation failed, which are also often easier for our neural network to get right. The fifth column shows the accuracy when excluding both empty fields and fields where our segmentation failed.