I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16270

Universal Investments in Toddler Health. Learning from a Large Government Trial

Jennifer L. Baker
Lise G. Bjerregaard
Christian M. Dahl
Torben S. D. Johansen
Emil N. Sørensen
Miriam Wüst

JUNE 2023

DISCUSSION PAPER SERIES

IZA DP No. 16270

Universal Investments in Toddler Health. Learning from a Large Government Trial

Jennifer L. Baker
Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital

Lise G. Bjerregaard
Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital

Christian M. Dahl
University of Southern Denmark

Torben S. D. Johansen
University of Southern Denmark

Emil N. Sørensen
University of Bristol

Miriam Wüst
University of Copenhagen and IZA

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Universal Investments in Toddler Health. Learning from a Large Government Trial*

Exploiting a 1960s government trial in Copenhagen, we study the long-run and inter-generational effects of preventive care for toddlers. We combine administrative data with handwritten nurse records to document universal treatment take-up and positive health effects for treated children over the life course. Beneficial health impacts are largest for disadvantaged children and may even extend to their offspring. While initial trial cohorts experienced positive health and socioeconomic impacts, those are absent for the final cohorts. This heterogeneity across individuals' background and cohorts documents that universal toddler care can alleviate inequalities at low costs, and that the counterfactual policy environment matters.

JEL Classification: I1, J1

Keywords: early-life investments, health, public policy, government trial, Denmark, digitization, automated transcription

Corresponding author:
Miriam Wüst
University of Copenhagen
Department of Economics
Øster Farimagsgade 5
1353 Copenhagen
Denmark
E-mail: miriam.w@econ.ku.dk

* We gratefully acknowledge financial support from the Independent Research Fund Denmark (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. Sørensen acknowledges support from the European Research Council (Starting Grant 851725). We thank the Copenhagen City Archive and Nicolai Eberholst for support with the archive material, and Jonas Færch Lotz for research assistance. We thank Andrew Barr, Meltem Daysal, Fane Groes, Jonas Cuzulan Hirani, Rasmus Landersø, Peter Sandholdt, Hans Henrik Sievertsen, and seminar participants at the University of Zurich, Uppsala University, Royal Holloway, Universidad Autonoma de Barcelona, Copenhagen Business School, Erasmus University Rotterdam, University of Southern Denmark, Stanford University, the Copenhagen Education Network, the NORC Demography Workshop (University of Chicago), the 2022 Essen Health Conference, the 2022 ESPE Conference, the 2022 NBER Summer Institute (Children Session), and the 2022 EALE Conference for helpful comments. In this project, we transcribe and use data from the "Copenhagen Infant Health Nurse Records" that was established by the Center for Clinical Research and Prevention, the Capital Region of Denmark, the University of Copenhagen, and the University of Southern Denmark. We thank Thorkild I.A. Sørensen for his contributions to the establishment of the CIHNR.

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 has been center stage in this work, more recently, the role of health policies, their timing, targeting, and content has gained interest in economic research. Motivated by models on early-life skill formation (Heckman, 2006; Attanasio, 2015), this work attempts to identify the role and relative importance of positive inputs in child health and human capital production.

Existing empirical research on the long-run importance of early-life health policies falls broadly into two streams: First, randomized trials with high-intensity model programs, such as the U.S. Nurse Family Partnership or the Perry Preschool Program, have made a strong case for 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, a literature exploiting large-scale administrative data and naturally occurring variation has documented positive long-run effects of access to early-life health policies on adult health and socioeconomic outcomes. Examples include studies on the impacts of nutrition and income support (Hoynes et al., 2016; Bailey et al., Forthcoming; Barr and Smith, 2023; Barr et al., 2022), health insurance and care (Wherry et al., 2018; Goodman-Bacon, 2018; Miller and Wherry, 2019; Noghanibehambari, 2022; East et al., 2023), early education programs (Rossin-Slater and Wüst, 2020; Bailey et al., 2021; Anders et al., 2023), and infant 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 use aggregate geographic and across-cohort variation for identification.

This paper combines the strengths of these two lines of empirical work: We exploit variation from a very large government randomized trial with preventive care for toddlers in Denmark in the 1960s. To study the impacts of this trial, we use individual-level information on program take-up from transcribed nurse records, and on childhood, long-run, and second generation outcomes from population administrative data. The nine-year trial allows us to study universal investments in toddler health (a thus far understudied childhood period), as well as heterogeneity of its impacts, and the dependence of impacts on the policy environment at large. These aspects are at the core of contemporary discussions on the design of universal vs targeted early-life health policies.

Specifically, we study a trial in the Danish capital that included all resident children born in 1959-1967. It extended an existing and well-functioning home visiting program for all infants during the first year of life to also cover the second and third years for children born on the first three days of each month. During the, on average, 12 to 14 first-year visits, nurses monitored infant health and development, counseled parents on investment decisions, such as infant nutrition and vaccination uptake, and referred ill infants to other health care providers, such as general practitioners (GPs). During the five second- and third-year visits, nurses continued to monitor toddler health, to encourage parental health investments (such as take-up of preventive care, vaccines and dental care, prevention of accidents, and provision of healthy nutrition), and to refer ill children to follow-up treatments. Additionally, they advised parents on less health-centered topics, such as the socio-emotional development of young children, parenting styles, and the take-up of childcare.

Our empirical work exploits the treatment assignment according to day of the month of birth, which effectively randomized children across all trial cohorts into treatment and control groups. We present both intention-to-treat (ITT) and instrumental variable results. To make our analyses feasible, we create and combine several data sources: Danish administrative data for the 1977-2020 period provides a range of health and socio-economic outcome measures for focal children of the 1959-1967 cohorts and their families.[1] A 1959-1961 cohort study allows us to measure childhood outcomes for three trial cohorts. Finally, to measure treatment assignment and background characteristics at the individual level (in the period pre-dating administrative data from Statistics Denmark), we transcribe handwritten nurse records for

[1] We measure focal children's and their siblings' mid-adulthood health and socio-economic outcomes (observed around the ages 30-60 years for the trial cohorts), as well as their mothers' completed fertility and labor market outcomes, and birth outcomes for their offspring.

the trial cohorts. 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; Dahl et al., 2023a) and scene, optical character, and handwritten text recognition (Goodfellow et al., 2013; Bluche et al., 2014; Lee and Osindero, 2016; Bluche et al., 2017; Bartz et al., 2021; Geetha et al., 2021; Kang et al., 2022; Dahl et al., 2023b).

We start our analyses by documenting a strong first stage and treatment effects for focal children. 57 percent of children born on the first three days of the month were assigned to the extended nurse program at age one. The share of non-compliance squares well with a reported 30 percent dropout rate in the universal first year program (Copenhagen Statistical Office, various years). Thus, exit from the first-year program mainly due to mobility resulted in children not being eligible for treatment assignment. Our strong first stage holds across birth years and months, and across family and child characteristics observed by nurses. This finding underscores that extended nurse care was a universal offer for resident children born on the relevant days of the month.

In the long run, treated individuals experience important health benefits. In our full trial sample, treated individuals score 4.4-5.2 percent of a standard deviation higher on a good health index that summarizes a set of hospital diagnoses and admission outcomes for individuals in their 30s through 50s (for improved statistical power). Examining underlying health measures for these relatively young and healthy individuals, we find improvements across the board for both diagnoses and hospital admissions, but only individually significant impacts for cancer and asthma (a seven and nine percent reduced probability of being diagnosed during our outcome data period). Moreover, treated individuals experience lower mortality in the short run–proxied as ever being observed in our outcome data in the absence of childhood mortality data–and the very long-run (a seven percent decrease in the probability of death before or in the year 2020 at the sample mean of 8.3 percent (ITT), or a 13 percent effect on treated individuals).[2] These results suggest that some program impacts manifest themselves

[2]As we discuss in detail, this large effect implies that our results from the main index analyses are likely to be lower bounds, as they ignore exit from our outcome data due to death.

at the oldest ages observed in our outcome data. Thus, the full health impact of the trial may be underestimated in our analyses.

While we find no impacts of extended home visits on a combined adult socio-economic status (SES) index in our main analysis, this finding conceals heterogeneity across gender and dimensions of the index: Treated women experience positive effects in labor market dimensions likely susceptible to adult health status (the share of time in employment during ages 30-50 and the probability of being out of the labor force at age 50). In line with stronger health results for females, these findings suggest positive impacts on labor market outcomes not primarily through a human capital but through a health channel in the very long run, when individuals in the work force make decisions about labor market exit.

The unique nature of the trial and our individual-level data on take-up allow us to assess the potential role of targeting of toddler care and the importance of a changing policy environment. First, today's policy debates, also in the Nordics, center around the impacts of universal vs targeted provision of care. Could providers such as nurses identify families that would likely reap the largest health benefits from toddler care in the family home? Exploiting our data on pre-treatment family characteristics, we show that the long-run health of treated children with initial health and social disadvantages (among them low birth weight children) are much more positively impacted by extended nurse care than average impacts indicate.[3] Moreover, for the offspring of focal children with initial disadvantages, we find positive impacts on health at birth, adding to the first order impacts.[4] While these results are more suggestive, they point to sustained benefits of extended nurse care in the second generation, given that poor health at birth is associated with large individual and societal costs (Currie and Moretti, 2007; Kreiner and Sievertsen, 2020).[5]

Second, the trial variation (across days of the month in all years) allows us to study the

[3]We show that heterogeneous impacts are not driven by a differential first stage. They may partly be due to more intensive treatment after randomization. However, we do not find evidence for a higher average number of second- and third-year visits for disadvantaged children.

[4]We find limited evidence for small fertility responses at the intensive margin.

[5]Our results are among the few demonstrating intergenerational effects of early-life investments in the health and development of children (Rossin-Slater and Wüst, 2020; Barr and Gibbs, 2022; East et al., 2023; Walker et al., 2023).

causal effects of universal toddler care across trial cohorts throughout the 1960s. While our analyses for the full trial sample show no impacts of extended nurse visiting on a combined SES index, we find positive and significant impacts on both health and SES, in particular labor market outcomes, for the earliest trial cohorts (born in the initial trial years 1959-1961). This heterogeneity may be due to three reasons, two of which (trial fidelity and spillovers in families) we can assess with individual-level data: The transcribed nurse records indicate that the fidelity of trial implementation was higher in the early years and we find suggestive evidence for spillovers in the family, especially from older (treated) children to their younger siblings. Those spillovers appear more important in the SES rather than the health domain. Given that more than one child of each family can be present in our trial data, spillovers likely attenuate our findings towards zero in the full sample. Two complementary analyses (one focusing on only the first-observed child in each family in the trial and one zooming in on a smaller sample of sibling pairs) support this reasoning and emphasize the need to carefully account for family outcomes in analyses of home-based interventions.[6]

A third factor related to the observed heterogeneity and likely instrumental for policy discussions today, is the role of a changing policy environment. Access to other influential policies in the final trial years, most prominently public childcare, made the extended nurse offer a less influential treatment. Even though the coverage rate with public childcare in Copenhagen was high already in the early 1960s (around 15 percent for the 0-6 age range), it doubled towards the end of the decade and thus the final trial years. This higher coverage of formal childcare for toddlers likely provided a similar treatment—with health monitoring, guidance to parents, and developmental and nutritional inputs—irrespective of trial status.[7]

Our final analysis of channels for the long-run impacts of the composite program points to exactly those elements as being influential. We rule out that fertility responses or labor

[6]By construction, children of the earliest trial cohorts are overly likely to be first-treated children in each family with multiple trial-exposed children. Sibling spillovers appear to be relevant for younger siblings but not older siblings. Younger, closely spaced siblings of treated children appear to be more likely to have a nurse record that is filled out and to be breastfed at one month if their older sibling is born on the first three days of the month. While these results are suggestive and based on constrained samples, they point to the importance of intensified nurse care during the first year for later life SES outcomes.

[7]Unfortunately, we cannot directly study childcare enrolment in the absence of individual-level data.

market decisions of mothers of focal children drive our results. We find that improved parental health investments (e.g., the uptake of additional preventive care and timely treatments) and improved childhood health are likely channels for the long-run impacts of the trial. Based on a 1959-1961 subsample of children, we document positive impacts on a good childhood health index parallel to our main analyses. This result supports our long-run results for asthma, which has been related to early childhood conditions, such as 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 results suggest that nurses were successful in preventing illnesses and promoting some health investments in the family, results for other dimensions of child development are less conclusive. This finding may be due to the quality of available measures, nurses' qualifications, or the dosage and content of the developmental advice given by nurses in the second- and third-year program.

By showing that universal toddler care has long-run consequences for treated children and families and that it can alleviate early life health inequalities, our study adds a central piece to the comprehensive literature on the short- and long-run impacts of early-life circumstances and health interventions such as nurse home visiting (Forsdahl, 1979; Barker, 1990; Currie and Almond, 2011; Almond et al., 2018; Wüst, 2022).[8] We show that universal counselling, information, and screening for families with toddlers has the potential to course-correct health trajectories of young children–not only infants–and may also provide benefits in non-health domains. When compared to existing work (predominantly on targeted policies and on policies in the in utero and first-year period), our results for the direct impacts of universal toddler care (provided in addition to universal first-year care) are smaller but economically relevant (especially for the early trial cohorts). Even extremely conservative accounts of the life-time health benefits created by the trial (focusing only on long-run mortality) indicate that its benefits outweighed its costs by a factor of 10. Importantly, given the large heterogeneity of long-run benefits across easily-observed dimensions of initial dis-

[8] Opposed to work based on roll-out variation, we observe individual-level treatment exposure and thus directly extend earlier studies on first-year universal preventive care (Wüst, 2012; Hjort et al., 2017; Bhalotra et al., 2017; Bütikofer et al., 2019; Rossin-Slater and Wüst, 2020).

advantage, our findings illustrate that universal childhood preventive care can help alleviate inequality even in a low-intensity and low-cost program as the one studied here.

Important for policy today, our study demonstrates that measuring impacts of interventions in the family home on the outcomes of others than focal individuals is instrumental for fully understanding their potential (Wüst, 2022). Moreover, families are rarely only exposed to one public policy and policy environments change. When the Copenhagen trial ended, national policy makers opted for a one-year nurse program, which became mandatory in all parts of the country in 1973. This decision was not based on analyses of trial data, but likely acknowledged the expansion of female labor force participation and the parallel increased childcare attendance of toddlers. Both factors made day-time home visits less feasible but also offered similar treatments. Thus our results speak to the important policy debate about the role of the counterfactual policy environment for the causal effects of early life policies. Our findings of larger positive health and SES impacts for trial cohorts with limited access to other influential treatments resonate with recent research on this topic (Kline and Walters, 2016) and highlight the importance of factoring in the interactions of policies.[9]

2 Institutional Background and the Copenhagen Trial

Universal home visiting for families with infants in Denmark dates back to 1937. The Danish National Board of Health (DNBH) centrally designed this preventive care program to combat high infant mortality rates of around six percent in the early 1930s. Exploiting the staggered introduction across municipalities in the 1937-1949 period, earlier research has documented both short- and long-run health benefits of program access (Wüst, 2012; Hjort et al., 2017).[10]

[9]Our results are in line with studies on the recent introduction of targeted nurse home visiting in the UK as an addition to existing family services, showing null effects for short-run outcomes (Robling et al., 2016). Similarly, earlier work suggests substitutability of access to home visiting and targeted preschool in Denmark for generating long-run impacts on health and SES outcomes (Rossin-Slater and Wüst, 2020).

[10]Original program guidelines mandated at least ten visits in the first year. The introduction of the program decreased infant mortality in the 1937-1949 period, saving 5-8 lives per 1000 live births. In the long run, exposed individuals were less likely to die and to be diagnosed with cardiovascular disease in the 45-64 years age range (a 1.3-2.8 percent decrease in the probability of being diagnosed at the control mean of 26.6 percent) (Wüst, 2012; Hjort et al., 2017).

Given improvements in living conditions throughout the 1950s, the DNBH planned to update program guidelines for the voluntary municipal program before making it a mandatory one. Infant mortality, in particular deaths from infectious diseases, had declined significantly to around two percent in 1960 (Det Statistiske Departement, 1964). In this low-infant mortality environment, the DNBH emphasized the need to focus on broader health monitoring and the encouragement of relevant parental health investments, both in the first year and during toddler years. The topics discussed bear similarity to elements in the early US Head Start program developed at the time (Barr and Gibbs, 2022), among them parental investments during toddler years, such as accident prevention, healthy nutrition, the uptake of dental care and toddler-year vaccinations. While relevant health services were in place (vaccinations) or expanding (free child dental care), documents from the time show that low parental uptake was a central concern (Copenhagen City Archives, various years). Additionally, the DNBH considered expanding the nurse treatment to cover non-health topics, including parenting styles and the socio-emotional development of children. This focus mirrored an increased focus on toddler development and interventions to support it in the family home.

In the early 1960s, interventions in the family home were particularly suitable. Midwife-assisted home births were the norm and predominantly pregnant women with identified health risks or social disadvantage (such as women who wanted to conceal their birth and opt for an adoption) were referred to hospitals.[11] Moreover, only one out of 10 children aged 0-6 was enrolled in formal childcare (Statistics Denmark, 2008). While childcare access was initially closely tied to social disadvantage as the central admission criterion (Rossin-Slater and Wüst, 2020), in parallel to large increases in the labor force participation of married women aged 25-29 to around 44 percent in the late 1960s, public childcare became an increasingly universal offer. Aggregate statistics presented in Appendix Table A.1 illustrate this development for Copenhagen: The coverage of public childcare for the under six year-olds increased from around 15 percent in 1959 to around 30 percent in 1970 in Copenhagen. As shown in earlier

[11] Hospital birth became the norm in Denmark in the 1970s (Vallgårda, 1996). Abortion was legalized in 1973.

work on public childcare, this expansion likely provided relevant overlapping developmental, health, and nutritional inputs for families (Rossin-Slater and Wüst, 2020; Attanasio et al., 2022). In our analyses, we discuss the relevance of this factor, which impacted increasing shares of children in both the treatment and control group across trial cohorts.

To inform policy decisions, the DNBH initiated a set of experiments with extended home visiting, among them the Copenhagen trial, which included children born between January 1, 1959 and April 1, 1968 (Copenhagen City Archives, various years; DNBH, 1970).[12] Its treatment group included all children born on the first three days of the month as well as all twins. Nurses offered the additional follow-up to families during one of the final first year visits, i.e., inclusion was conditional on Copenhagen residence and participation in the first-year program.[13] Importantly, the Copenhagen trial was conducted largely without additional funding and manpower. To compensate nurses for the additional visits to 10 percent of, on average, 160 children in each nurse district, they were instructed to reduce the number of first year visits for all families by around one to two visit during the trial period (from the pre-trial average of around 14 (Copenhagen City Archives, various years)).

Based on nurse records and instruction material from the Copenhagen City Archives, Appendix Tables A.2 and A.3 detail the content of the first-year and the extended program visits of the trial: During the first year, nurses monitored infant health and development,

[12]None of the other projects across Denmark used randomization. Municipalities in the counties of Holbæk (from 1960), Esbjerg (from 1961), and Aarhus (from 1962) extended nurse care beyond the first year of life either for all families or subsets of families with health and social risk factors. We have no preserved nurse records from these areas.

[13]The Copenhagen program head, Dr. Biering-Sørensen, played a central role in the organization of the trial, as documented in archive material (Copenhagen City Archives, various years): First, he had previously conducted several initiatives, including a WHO-sponsored examination of side effects of the Calmette vaccine in infants, a project on pregnancy visits for selected mothers, and a project on documenting the prevalence of first-year health issues in a sample of children born on the first three days of each month. Thus, the Copenhagen program was geared to implementing new program elements. Second, Biering-Sørensen focused on the quality of nurse registrations and on including structured data that would be easy to use in research. Thus, from the onset, the trial used suitable nurse records to be able to follow up on its results. Third, Biering-Sørensen's focus on scientific methods (choosing a random set of families as treatment group and focusing on twins to potentially compare outcomes of mono-zygotic and di-zygotic twins) was visionary, even though power calculations were not part of the design. Additionally, archive material documents extensive preparation meetings prior to and regular staff meetings during the trial on practical aspects and the fidelity of implementation. Finally, archive material illustrates that the trial was high on the political agenda, with Copenhagen decision makers, including deputy mayor for social affairs Urban Hansen, explicitly supporting and approving the trial.

nutrition, and family living conditions, including measures of mothers' health and labor force participation. At around age 12 months, nurses recorded information on vaccination uptake and health care usage. The five second- and third-year visits (around child age 15, 18, 24, 30, and 36 months) continued to have a strong focus on health and the early identification of illnesses. In an environment with few scheduled interactions with professionals during toddler years, nurses encouraged uptake of relevant care. They educated parents about the prevention of accidents and the benefits of investments, such as vaccinations, dental care and teeth brushing, and healthy nutrition. Moreover, nurses monitored the development of children in non-health domains and advised parents on parenting style and childcare attendance.

3 Data Sources

We combine three main data sources: First, we transcribe handwritten nurse records for the 1959-1967 cohorts from the Copenhagen City Archive (Bjerregaard et al., Forthcoming).[14] Second, we use administrative data on outcomes at Statistics Denmark for focal children, their siblings, parents, and own children. Third, we use data on childhood health and development from the Copenhagen Perinatal Cohort (CPC) covering children born in 1959-1961 in Rigshospitalet, the largest hospital in Copenhagen (Zachau-Christiansen and Ross, 1975).

3.1 The Copenhagen Nurse Records: Transcription and Linkage

Data Extraction from Handwritten Nurse Records Figure 1 depicts a scan of a multi-page nurse record for a treated infant. The record, while hand-written, has a standardized layout: The first page contains a table for structured nurse observations at specific ages during first year visits. Registrations are structured around the child's age (rows) and across different topics (columns). As there were typically more visits than recorded in this table, the second page (the flip side) contains space for information on date, child weight, and less

[14] The archived data only includes the cohorts 1959-1967 but no records for pre- or post-trial years. The time frame of the preserved records suggests that the records were archived due to the trial.

structured observations in free text. Finally, the third page contains a table on topics (rows) for the second- and third-year visits (columns). The presence of this "treatment table" in a record indicates that the child was offered the extended program.[15]

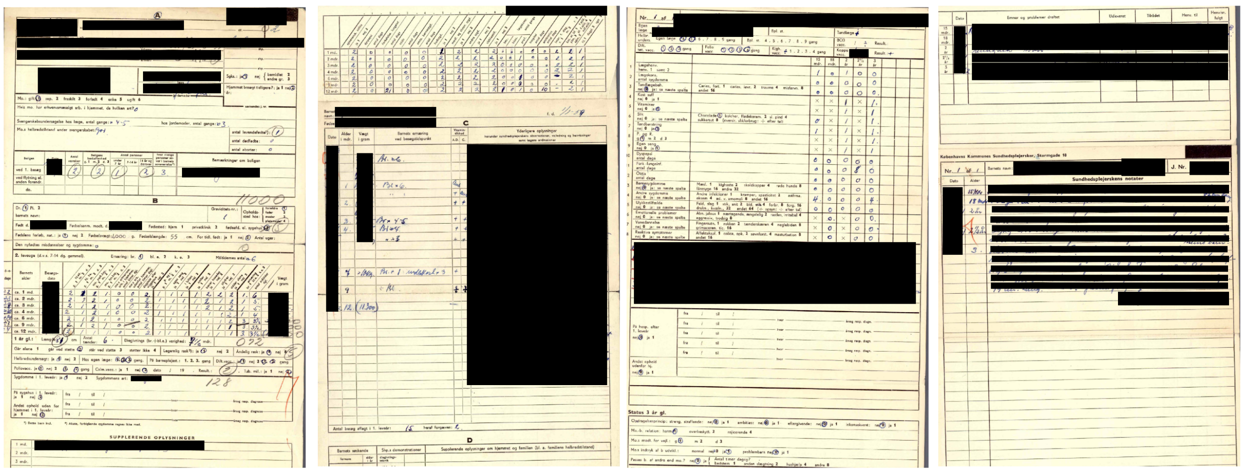


Fig. 1 Sample Nurse Record for a Copenhagen Child.

Notes: The pages depict a scanned nurse record of a child. 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 primarily for nurse comments in free text. 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.

Manual transcription of the nurse records is not feasible due to the large number of records and fields. We instead apply a machine learning approach to transform the images of the scanned records into structured data for our analyses (for details, see Appendix E). First, we use an unsupervised layout classification approach based on clustering to identify the presence of the treatment table.[16] Second, we use a neural network to classify whether each cell of the treatment table is filled out or not. Third, we use neural networks to transcribe fields of the nurse records containing registrations for first-year visits. We train the networks

[15]The treatment table is not always the third page of the record. Record length varies (from 1 to 20 pages), with the vast majority of records having 2 or 4 pages. Instruction material preserved in the City Archive states that nurses were to add additional pages to the record of treatment children when offering the extended nurse program in one of the final first-year visits. Browsing the paper records, we have established that nurses typically stapled the page with the treatment table to the original record.

[16]Our approach does not require labelled training data. However, we manually classify 4,000 nurse records to assess the performance of our clustering method. We find that our treatment table predictor has precision and recall of unity.

on a manually transcribed subset of the records.[17] In Appendix E.2, we show that the neural networks transcribe fields with an accuracy in the 95-99 percent range for most table fields, which is on par with the accuracy of crowd-sourced transcriptions in Dahl et al. (2023a).[18]

Linkage and Coverage of Nurse Records We link the nurse records to the other data sources in three steps (illustrated in a flowchart in Appendix Figure A.1): First, we start with the full set of scanned paper records to identify 92,902 relevant nurse records.[19] Second, we use manually transcribed information on mothers', fathers', and children's names and dates of birth to link the records to the Danish Central Person Registry (CPR), containing a unique personal identifier for each individual residing in Denmark in any year from 1968 on. The matching algorithm provides a CPR-link for 88,808 (96 percent) of the relevant records, while the remaining 4,094 (four percent) do not result in a match.[20] Third, we merge the nurse records to administrative data on outcomes at Statistics Denmark using the unique personal identifier. 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, children, and first spouses of the focal individuals, if ever observed in the administrative data. We merge all but 808 nurse records with an identifier to data at Statistics Denmark. Given that those individuals were Danish residents in 1968, we document that the majority of non-linked records refer to individuals who left Denmark in the 1968-1977 period.[21] In sum, 4,902 relevant records (5 percent) remain unmatched to the

[17]Most of the manual transcriptions were done in previous works by Andersen et al. (2012) and Bjerregaard et al. (2014), leaving minimal need for additional labelling

[18]As our transcription methods rely on segmentation of field-level images, our method fails if the quality of a scan is particularly poor and segmentation is not possible. This loss depends on factors such as the quality of the scans and the condition of the source document, which are likely independent of treatment assignment.

[19]Among the scanned paper records, we have manually identified and excluded duplicate scans and other types of documents (e.g., notice of a family moving) that have been archived together with nurse records. 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 in the nurse archive.

[20]Failure of the linkage can be due to three reasons: Incomplete records (missing names and birth dates), death or emigration of individuals prior to 1968, or poor quality of the handwritten information on names and dates of birth. We analyze the potential impact of selective emigration or mortality in Section 5.3.

[21]The CPR provides a yearly updated residency status for all individuals ever registered as residents since 1968. This status variable is not available for the pre-1977 period in the data at Statistics Denmark. We have obtained it directly from the CPR office. We cannot rule out that individuals who emigrated died abroad.

administrative outcome data.[22] A final exclusion in our analyses relates to multiple births: As all twins were offered the extended program, 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, leaving us with 87,336 individuals.

Do the preserved nurse records provide good coverage of the children in the Copenhagen nurse program, and was the first-year program universally accessible for all Copenhagen residents? Appendix F uses aggregate statistics to document that the nurse records cover the vast majority of infants entering nurse care in Copenhagen, and that the nurse program had universal outreach among Copenhagen residents. The aggregate data also report a discontinuation rate in the first-year nurse program of around 30 percent (indicating that non-compliance in the trial is strongly related to dropout of infants prior to treatment assignment).

3.2 Administrative Data from Statistics Denmark

To study human capital, labor market status, and health outcomes, we combine data from 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).[23] To confront multiple hypotheses testing and to increase statistical power, we create indices for socioeconomic (SES) and health status in adulthood. To form the combined 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,[24] and the share of time in employment during ages 30-50. Our good health index aggregates inpatient contacts with hospitals and diagnoses given in hospitals related to

Thus, the CPR status may understate deaths in the non-merged group.

[22]Appendix Figure A.2 suggests that the share of unlinked nurse records due to a missing identifier or missing outcome data at Statistics Denmark is stable across cohorts, with slightly higher success rates for younger cohorts. Across the treatment and control groups, match rates are slightly higher for treatment days. We return to and directly study this pattern in our analyses of death and migration outcomes as higher linkages for treated days-of-the-month may indicate program impacts on early mortality.

[23]As explained, to enter our analysis samples for health and SES outcomes, individuals have to be observed in the register data at Statistics Denmark after 1977/1980. However, for all individuals in the nurse records with a valid personal identifier in 1968, we can track pre-1977 deaths and emigrations.

[24]We winsorize earnings (one percent of both tails for each age and after adjusting earnings for inflation).

early-life circumstances. Specifically, it includes 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 (all in the 1977-2018 period): Diabetes, cardiovascular disease, cancer, asthma, any mental health issues, or infections.[25]

Following Kling et al. (2007), we create our indices as equally weighted averages of the z-scores of the non-missing underlying variables, orienting the signs such that a more beneficial score is associated with a higher value. We calculate these z-scores by subtracting the means and dividing by the standard deviations of the control group, i.e., individuals *not* born in the first three days of a month. Alternative weighting schemes for the good health index (putting relatively more weight on the hospital admission measures as broader proxies of health rather than the individual diagnoses) and exclusion of specific subsets of diagnoses do not impact our conclusions. For interpretation purpose, we rescale our indices by dividing by the standard deviation of the control group. This step ensures unit standard deviation of our indices but has no impact on the statistical significance of our results.

For both our SES and health index that summarize outcomes over periods of adulthood (over 20 and 29 years), we have to consider missing data issues. Individuals who die or leave the country are not observed in the outcome data. For the SES index, the outcome is missing for individuals with missing values for all outcome years for either the underlying education or labor market outcomes.[26] For the health index, we set both admissions and diagnoses to zero for individual not observed in the hospital data in a given year, resulting in a larger sample for this outcome. This strategy is the conservative choice because it likely overstates the health of individuals who either died or emigrate (and instead of missing are set to zero, i.e., having no diagnosis). Importantly, we show that our handling of missing data and potential selective attrition do not drive our results: First, we perform our analyses on

[25] Appendix H presents the ICD codes used. We cannot exclude females' birth-related hospitalizations consistently over time. While birth hospitalization length has been declining over time, this decline should be differential by day of the month of birth for our estimates to be biased. We find no impacts on timing of fertility (any children, number of children, and age at first birth) for treated females.

[26] If individuals are observed in a subset of years, we generate average earnings and employment outcomes based on those. Between 2-4 percent of our sample have missing values for educational attainment or labor market outcomes, which leads to around five percent of individuals with a missing SES index.

the sub-sample of individuals who survive and live in Denmark in 2017. Second, we directly analyze the impact of being born in the first three days of the month 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. In these analyses we use (i) our main sample of matched nurse records and, alternatively, (ii) all 92,279 records. For unmatched records, we impute individuals' status as either being "dead" or "having emigrated".

The analyses based on imputed samples help us assess selection out of our data due to emigration and deaths, but they are also interesting in their own right. Importantly, as mortality is still relatively low for the age groups that we study (and given that the vast majority of deaths occur early in life or during the final years of our outcome data), we focus on a relatively coarse outcome (probability of death) rather than other types of survival analysis. We show complementary descriptive evidence for the impacts on the timing of mortality using non-parametric survival analysis.

Finally, as the administrative data allow for a family link, we study outcomes of other family members of focal children: siblings, biological mothers, and focal children's own offspring. First, as nurses visited family homes, spillovers from trial children with second- and third-year visits to siblings may occur. These spillovers in the family may attenuate our findings but are also a relevant outcome (potentially constituting a benefit of this program provided in the family home). We study a sample of trial children and all singleton children born to the same mother in the 1949-1977 period, a total of 131,700 children. We zoom in on the first two children in each family (80 percent of those children). Their median spacing is three years. We create separate sibling samples with either focal first-born or later-born children and their immediate younger or older sibling, respectively.[27] Second, for mothers of trial children, we consider subsequent fertility, focusing on first-time mothers observed in our trial data (age at birth, spacing between births, and total number of children), as well as labor market outcomes (average income at ages 40-65 and the share of years the mother is

[27]We only consider families with a focal child who has a nurse record and thus was likely a trial-eligible Copenhagen resident. These two samples consist either of siblings who have an older trial-exposed child in the family (N = 35,637) or of siblings with a younger trial-exposed child in the family (N = 28,349).

employed during ages 40-65). These analyses may also shed light on potential mechanisms for impacts on treated children. Third, for the offspring of focal individuals, we study health at birth, measured as birth weight in grams, low birth weight status, and preterm birth status (birth prior to 37 full weeks).[28]

3.3 The Copenhagen Perinatal Cohort 1959-1961

To study intermediate childhood impacts of the trial, we use data from the Copenhagen Perinatal Cohort (CPC). This cohort study follows children from three cohorts from birth through follow-up mother surveys and examinations at their hospital of birth (for details on content and variable construction, see Appendix G).[29] To exclude non-Copenhagen residents from our analyses, we only consider singleton children in the CPC who also have a nurse record (N = 4,369). Given attrition, we use samples of between 1,800 and 2,700 children with data from the three and six year follow-up. While sample sizes vary across outcomes, attrition should be systematically related to the day of the month of birth for our analyses to be biased. As we show, being born on the first three days of the month does not predict participation in the three year data collection (our main focus).[30]

Our main childhood outcomes are indices for good child health at age three and six and an index for overall child development at age three. We construct them to increase statistical power using the same approach as in our main analyses. Additionally, we separately study CPC measures, as detailed in Appendix G: First, we construct indicators for the child having been hospitalized, having complied with the vaccination schedule, having been exposed to infectious diseases, and having consumed antibiotics by ages three and six. We interpret antibiotics consumption as measuring disease exposure given large-scale and easy access to

[28]We discuss first-generation fertility impacts, which are minor. We have also considered educational attainment for second-generation children born prior to 1995 and thus old enough to be observed at appropriate ages. The sample for these supplementary analyses contains around 64 percent of the full second-generation sample and we study the probability of having completed more than compulsory education.

[29]The CPC over-represents mothers with medical issues and low SES status (Schack-Nielsen et al., 2010).

[30]We do not merge the CPC data with the transcribed data from the nurse records at Statistics Denmark (but only with information on having a nurse record). Thus, we perform reduced form analyses using the CPC and scale our results with the first stage from our full sample analyses.

the drugs (with around 48 percent of children having ever consumed antibiotics at the given ages). Exposure to childhood disease has been linked to adult health outcomes (see, for example, Crimmins and Finch, 2006). Second, as a measure that is responsive to childhood health and nutritional investments, we consider height during childhood.[31] Third, we study a set of six separate child developmental milestones indices, which together make up our child development index. They combine information on the age at completion for a set of 20 tasks across six dimensions (such as motor or language development) reported at the three year follow-up. We create the indices for these six domains of child development as in Flensborg-Madsen and Mortensen (2018).

4 Empirical Methods

All our analyses compare average outcomes of trial children and their families across birth day of the month groups. We regress our outcome Y_i for individual i on Z_i, an indicator for whether individual i was born in the first three days of a month. We identify the intention-to-treat effect of being born on these days and also provide instrumental variable estimates for the impact on compliers. To increase precision (and to informally assess the validity of our design), we estimate specifications with year and month of birth fixed effects, γ_{YOB} and λ_{MOB} (accounting for systematic differences in outcomes across birth cohorts and seasons), day of the week of birth fixed effects, δ_{DOW} (accounting for potential differences in timely access to the nurse program), and family characteristics observed by nurses, X_i:

$$Y_i = \alpha + \beta Z_i + \gamma_{YOB} + \lambda_{MOB} + \delta_{DOW} + \theta' X_i + \varepsilon_i.$$

Individual-level characteristics at birth include an indicator for the child's low birth weight status (birth weight below 2,500 grams, transcribed from the nurse records) and indicators for

[31] Height measures at three and six years display large shares of missing values (due to physicians not reporting the actual height of the child but only ticking a box for a height range, the latter case resulting in a missing value for height). This feature of the data limits our analyses and prevents us from including height in the good health index.

child sex, whether the child was firstborn, whether the child was born in Copenhagen, whether the father-registration is missing in our data, whether the child was born in a hospital, and the mother's age at birth, which are all observed in the administrative data.[32]

We perform heterogeneity analyses across cohorts, child sex, low birth weight status, and a composite measure of initial health and social disadvantage. These two dimensions of disadvantage are hard to separate in their impact on children's health and likely jointly considered by nurses. Thus we define an initial disadvantage indicator as being one 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 (a marker of health or social disadvantage at the time), or had a missing father registration.[33]

In all our analyses, we rely on the pragmatic randomization in the trial and assume that being born between the first and the third of a month does not have an impact on outcomes of assigned children and families through other channels than the extended nurse program. The program extension can impact families both directly and through down-stream treatments and parental behavioral responses, all of which will be captured in our estimates for the impact of the trial. There are two main threats to identification: First, being born between the first and the third of a month could affect child outcomes directly and through other channels, for example, if families have more resources at the beginning of a month (a payday effect). To assess the potential importance of this factor, we constrain our control group and locally compare children born on the first three days of a month to those born in the four days after. Moreover, we run our analyses on a sample of children from the same cohorts but born in the three biggest Danish towns outside Copenhagen: Aarhus, Odense, and Aalborg. In the absence of extended nurse care assigned in a similar way as in Copenhagen, we do not

[32]As we observe mothers' highest obtained educational attainment only in the administrative register data after 1980, it is measured retrospectively. Including it in our main analyses does not alter our conclusions.

[33]Combining alternative measures of social and health disadvantages does not alter our conclusions. In supplementary analyses we consider heterogeneity across a broad range of individual measures (whether the mother was above or below median age at birth in our sample, birth location (home or hospital), whether we observe the father in administrative data, whether the mother has completed more than compulsory education, breastfeeding status at one month, and child parity (whether the child is firstborn or of higher parity).

expect to see any effects of day of the month of birth for these individuals.[34]

Second, the assignment of treatment to one individual may affect the outcomes of other individuals. These spillovers, most importantly within families, may result from parents learning and applying information to all children in the family or nurses also directly focusing on them once in the family home. As described earlier, we analyze the importance of spillovers in samples that rule them out and in sibling samples that zoom in on them.

5 Results

5.1 Descriptive Statistics

Table 1 presents means and standard deviations for birth, family, and first year characteristics from the nurse records (Panel A), background characteristics from the administrative data (Panel B), and adult outcomes (Panel C) for our main analysis sample and samples defined by the child's day of the month of birth.[35] The final column presents the p-value from a t-test of equality of means across children grouped by their day of the month of birth.

The top panels confirm that children are very similar across the groups. While some differences are statistically significant, those are economically small and do not indicate large differences across groups.[36] Around half of the children in our sample are firstborn children and 22 percent of children were born in a hospital. These figures distinguish our sample from the population outside Copenhagen.[37] As is apparent in the bottom of the table, we consider

[34]Our administrative data (and the nurse records) only include information for the cohorts 1959-1967. Thus, we cannot use pre- or post-trial cohorts in Copenhagen as natural placebo groups. In Aarhus, the municipality experimented with longer follow-up of at-risk families from 1962. This project selected families according to risk rather than the child's date of birth.

[35]Appendix Figure A.3 confirms that the number of nurse records does not vary systematically across days of the month of birth.

[36]A joint test of the predictive power of all first-year characteristics for determining a child's day of the month of birth-group is unfeasible due to small sample size when conditioning on the non-missingnes of all the shown covariates. Focusing on the set of characteristics measured by nurses at one month (or observed in the administrative data), we cannot reject that they are jointly zero and thus not predict treatment group membership (with a p-value 0.36).

[37]Appendix Table A.5 shows that in the 1959-1967 period, capital births relative to births in the rest of the country are to younger parent, are more likely to be firstborn, and are more likely to take place in a hospital.

Table 1 Descriptive Statistics (means, std.dev.).

	Full Sample	Born 4-31	Born 1-3	N	p-value
A. Background Characteristics, Nurse Records					
Birth prior to due date (%)	11.93 (32.41)	11.92 (32.40)	12.01 (32.51)	83,093	0.81
Low BW (%)	4.97 (21.73)	4.98 (21.76)	4.83 (21.44)	87,115	0.54
Weight at birth	3327.60 (527.12)	3327.84 (528.22)	3325.38 (516.75)	87,115	0.68
Weight at 1 mo.	4057.60 (622.47)	4061.63 (623.33)	4022.43 (613.79)	80,784	0.00
Weight at 2 mo.	5022.04 (741.89)	5021.51 (721.51)	5026.55 (896.17)	77,663	0.56
Weight at 3 mo.	5852.53 (840.42)	5860.35 (810.70)	5787.56 (1053.26)	76,061	0.00
Weight at 4 mo.	6670.62 (1037.38)	6677.72 (1013.45)	6611.92 (1215.92)	76,146	0.00
Weight at 6 mo.	7935.70 (1079.44)	7947.19 (1046.44)	7844.84 (1308.13)	72,924	0.00
Good/avg. home economic status at 1 mo. (%)	93.64 (24.41)	93.59 (24.48)	94.01 (23.74)	55,042	0.24
Good/avg. cleanliness at 1 mo. (%)	99.87 (3.59)	99.86 (3.68)	99.93 (2.60)	76,078	0.12
Lifts head at 2 mo. (%)	51.57 (49.98)	52.18 (49.95)	46.12 (49.85)	62,897	0.00
Lifts head at 4 mo. (%)	98.34 (12.79)	98.30 (12.92)	98.63 (11.61)	55,743	0.07
Babbles at 2 mo. (%)	83.80 (36.85)	83.98 (36.68)	82.13 (38.31)	71,345	0.00
Babbles at 6 mo. (%)	99.79 (4.56)	99.80 (4.48)	99.73 (5.20)	44,112	0.31
Sits at 6 mo. (%)	23.51 (42.40)	23.43 (42.36)	24.15 (42.80)	56,256	0.22
Sits at 9 mo. (%)	92.77 (25.90)	92.80 (25.85)	92.47 (26.40)	60,302	0.34
Exclusively breastfed at 1 mo. (%)	57.44 (49.44)	57.43 (49.45)	57.60 (49.42)	77,776	0.78
Exclusively breastfed at 2 mo. (%)	33.30 (47.13)	33.32 (47.14)	33.05 (47.04)	74,626	0.63
Exclusively breastfed at 3 mo. (%)	20.80 (40.59)	20.81 (40.60)	20.64 (40.48)	72,771	0.74
Exclusively breastfed at 6 mo. (%)	3.07 (17.25)	3.06 (17.22)	3.16 (17.50)	70,271	0.65
Childcare attendance at 6 mo. (%)	7.85 (26.90)	7.64 (26.56)	9.91 (29.88)	60,147	0.00
Childcare attendance at 12 mo. (%)	8.53 (27.94)	8.44 (27.79)	9.40 (29.18)	59,274	0.01
Good/avg. mother mental health at 1 mo. (%)	99.14 (9.24)	99.12 (9.32)	99.28 (8.46)	56,162	0.23
Good/avg. mother mental health at 12 mo. (%)	98.77 (11.01)	98.76 (11.06)	98.87 (10.57)	55,035	0.49
Good/avg. mother physical health at 1 mo. (%)	99.40 (7.71)	99.39 (7.77)	99.50 (7.09)	58,081	0.34
Good/avg. mother physical health at 12 mo. (%)	99.38 (7.87)	99.37 (7.91)	99.44 (7.44)	57,618	0.50
All Cells Empty, 9 mo. (%)	24.02 (42.72)	24.00 (42.71)	24.13 (42.79)	87,119	0.79
Treatment table (%)	5.86 (23.50)	0.26 (5.07)	57.03 (49.51)	87,336	0.00
B. Background Characteristics, Admin. Data					
Female (%)	48.88 (49.99)	48.82 (49.99)	49.44 (50.00)	87,336	0.27
Born in CPH (%)	81.74 (38.63)	81.74 (38.63)	81.76 (38.62)	86,944	0.96
Hosp. birth (%)	21.87 (41.34)	21.92 (41.37)	21.41 (41.02)	87,336	0.28
Mother's age at birth	24.98 (5.37)	24.98 (5.37)	24.98 (5.36)	86,754	0.97
First-born Child (%)	55.20 (49.73)	55.19 (49.73)	55.26 (49.73)	86,754	0.91
Father missing (%)	5.43 (22.67)	5.41 (22.62)	5.66 (23.10)	87,336	0.33
First child of the family in the trial (%)	74.47 (43.60)	74.48 (43.60)	74.31 (43.69)	86,754	0.73
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	0.68
Above compulsory educ. (%)	74.09 (43.82)	74.07 (43.83)	74.25 (43.73)	85,330	0.72
Higher education (%)	29.11 (45.43)	29.13 (45.43)	28.98 (45.37)	85,330	0.78
Earnings at 25 (DKK 1000)	190.54 (133.81)	190.57 (133.80)	190.30 (133.93)	85,741	0.86
Avg. earnings 30-50 (DKK 1000)	286.50 (181.80)	286.31 (181.92)	288.27 (180.71)	85,035	0.35
Avg. empl. age 30-50 (%)	80.14 (29.00)	80.08 (29.07)	80.69 (28.37)	84,100	0.07
Out of labor force, age 50 (%)	16.97 (37.53)	17.04 (37.60)	16.34 (36.97)	78,595	0.12
Diabetes (%)	4.08 (19.79)	4.11 (19.86)	3.81 (19.15)	87,336	0.19
Cardiovascular disease (%)	26.74 (44.26)	26.79 (44.29)	26.27 (44.01)	87,336	0.30
Heart disease (%)	5.33 (22.47)	5.35 (22.51)	5.18 (22.17)	87,336	0.51
Asthma (%)	5.24 (22.29)	5.29 (22.39)	4.79 (21.35)	87,336	0.05
Cancer (%)	7.72 (26.69)	7.78 (26.78)	7.18 (25.81)	87,336	0.05
Any mental health contact (%)	14.12 (34.82)	14.13 (34.83)	14.02 (34.72)	87,336	0.77
Infection (%)	18.90 (39.15)	18.92 (39.16)	18.77 (39.05)	87,336	0.74
Hosp. nights 30-39	6.12 (19.45)	6.14 (19.63)	5.92 (17.70)	87,336	0.30
Hosp. nights 40-49	4.77 (19.20)	4.81 (19.15)	4.47 (19.65)	87,336	0.12
Hosp. nights 50-59	3.61 (16.13)	3.62 (16.04)	3.49 (16.85)	87,336	0.48
Ever hospitalized (30-59) (%)	75.89 (42.78)	75.95 (42.74)	75.30 (43.13)	87,336	0.18

Notes: Earnings are in DKK1,000 (2015 DKK). The diagnoses are indicators that equal one if an individual has ever been diagnosed with the condition in our outcome data. The p-value in the final column is from a t-test for equality of means between the day of the month of birth-groups.

a sample with relatively low disease prevalence. This fact motivates our index construction.

Appendix Table A.4 presents descriptive statistics for the CPC sample with pre-determined characteristics in Panel A and childhood outcomes in Panel B. We observe few differences across groups. Importantly, day of the month of birth is not related to whether the child is observed in the three-year examination. As an exception, significantly fewer children are female and are born with a low birth weight between the first and the third day of the month.[38] For outcome measures, the table tentatively suggests that children born on the first three days of the month reach developmental milestones earlier and are less likely to have been treated with antibiotics. Children born on the first three days of the month are taller at ages three and six, but the samples for these outcomes are much smaller and we are careful in interpreting this difference. We more formally assess longer-run and childhood impacts of the extended nurse treatment in Section 5.3.

5.2 First Stage: The Assignment of Extended Nurse Visiting

Figure 2 shows that we classify 57 percent of nurse records for children born between the first and the third of a month as containing a treatment table (our first stage). This share is stable across birth months, quarters, and years, indicating that the pragmatic trial was in place over the entire data period. We also find that the share of treatment table records is very similar across geographic units defined by groups of parishes within the municipality of Copenhagen (Appendix Table B.1), supporting that the trial covered all of Copenhagen.

While our main measure of treatment assignment is the presence of a treatment table in a child's record, Figure 3 documents the intensity of the treatment that was administered, i.e., trial fidelity. Between 50 and 95 percent of the treatment tables have at least one cell filled out at the relevant child age, with higher shares of treatment tables with at least one registration for the earliest visits in the second year. For the cohorts born in the final years of the trial, there are fewer registrations in the child's third year of life potentially indicating a lower

[38] This imbalance is not present in the full CPC sample but emerges in our analysis sample of children who are observed at three years. We address this imbalance by including low birth weight status as a control variable and by exploring results in separate samples of females and males.

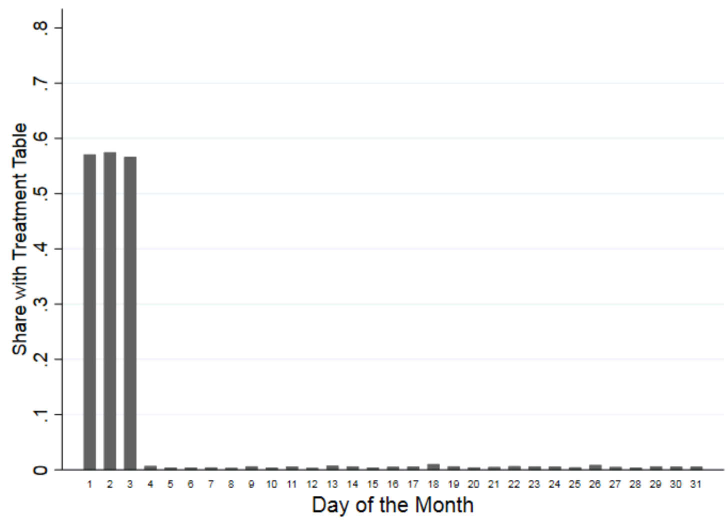


Fig. 2 Share of Nurse Records with a Treatment Table by Day of the Month of the Child's Birth.

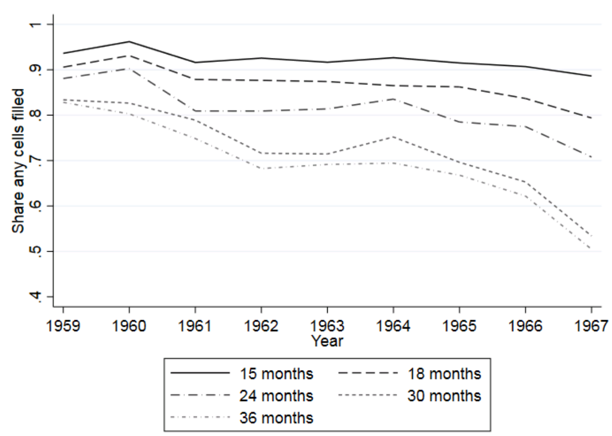


Fig. 3 Completeness of Nurse Registrations in Treatment Tables across Visits and Birth Cohorts.

Notes: In this figure, we classify the cell entries in the extended program for the visits at ages 15, 18, 24, 30, and 36 months and measure whether at least one cell was filled out at the visit. The figure shows the share of treatment table records where at least one cell for an age-specific visit is filled out, by year of birth of the child.

fidelity of implementation of third-year visits in the final trial years. While we explore this variation as one dimension of heterogeneity that can help inform us about the effectiveness of different elements in the nurse program (partial vs full trial program exposure), in general, we conclude that the presence of a treatment table is a good measure for the visits having taken place. Focusing on the average number of visits, we find that children born on the first three days of the month receive 2.6 and 1.4 visits out of the scheduled three and two visits in the second and third years of life, respectively.

Panel A in Appendix Table B.2 presents regression equivalents for the graphical evidence on the first stage. It holds across specifications and when restricting our sample to only those born between the first and the seventh of a month.[39] Appendix Table B.3 probes the robustness of our first stage to changes of sample: Panel (A) confirms that it holds when we include the non-merged records (excluded in our main analyses). For this sample, we add a set of controls for factors observed by nurses during the first month of the child's life, all transcribed from the nurse records.[40] Panel (B) shows that the first stage is robust in our main analysis sample when we only select nurse records that we confidently classify as Copenhagen residents (proxied as them having at least one nurse registration in their record at 12 months). 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, we find a strong and slightly larger first stage estimate of around 57-65 percent. As this sample is selected on our ability to transcribe the individual nurse records and data on first-year registrations, we prefer our full sample of merged nurse records for the main analyses.

Finally, given our data on child and family characteristics measured prior to the initial nurse contact, we perform a complier analysis. This analysis may also reveal whether nurses

[39]The remaining panels in Appendix Table B.2 show the 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).

[40]Appendix Figure B.1 illustrates this point by match status of the records. For individuals observed in 1968 with a personal identifier, we document a very similar pattern of treatment table records across the days of the month. Those without an assigned identifier in 1968 have a much lower share of treatment tables. These records are both more likely to be of poor quality, complicating transcription, and 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.

targeted certain types of children and families in the trial even though this was explicitly not intended. As we illustrate in Appendix Table B.4, using our first stage specification with year, month, and day of the week of birth fixed effects, we observe a strong first stage across all groups. This result suggests good compliance with universal assignment. Children born in hospitals and children with a missing father registration have the smallest first stage estimates (with considerable overlap between the two groups). We attribute this fact to the special pattern in hospital births at the time, with high shares of mothers with social and health disadvantage. Thus, we expect larger first-year mobility and mortality among hospital births and thus more drop-outs from the nurse program prior to treatment assignment.

5.3 Main Results

Long-run Impacts of Extended Home Visiting on Health and SES Outcomes Table 2 presents reduced form and instrumental variable estimates for the SES and health indices using our full sample of merged nurse records. Mirroring the descriptive statistics, there is no significant effect of being born on the first three days of the month on the SES index and estimates are small. The reduced form results suggest a robust and significant positive effect on our health index of 0.026-0.029 standard deviations, translating into instrumental variable results of a 0.044-0.052 standard deviation increase. As illustrated in Appendix Table C.1, there is no similar health impact of being born on the first three days of the month in Aarhus, Odense, and Aalborg, the three largest Danish towns outside Copenhagen. This analysis directly rules out day of the month effects.[41]

While these main results suggest long-run impacts on our health but not SES index, they conceal heterogeneity across gender and underlying outcome measures.[42] First, we find

[41]We select individuals born in 1959-1967 with a birth parish in one of the three towns. We cannot control for birth weight in the administrative data. Addressing remaining concerns about spurious effects, we perform a permutation test. We randomly assign the day of the month of birth to individuals in our Copenhagen sample 1,000 times. Plotting the cumulative distributions of t-values, the true value for the health but not the SES index regressions falls in the tail.

[42]We conduct a simple power analysis calculating the ITT effects that we should be able to detect. Appendix Table C.2 shows the required effect sizes. For example, an effect on average income at ages 30-50 would have to be no greater than DKK 5,859 (in 2015 DKK), i.e., an increase of around two percent of the mean of the dependent variable, for us to be able to detect it.

Table 2 Main Results: The Effect of Extended Nurse Care on SES and Health Indices, ITT and IV Estimates.

	(1)	(2)	(3)	(4)
Panel A: SES Index – ITT				
Born 1-3	0.007	0.006	0.001	0.005
	(0.012)	(0.012)	(0.011)	(0.014)
No. of obs.	83,029	83,029	82,391	18,868
Panel B: SES Index – LATE				
Treatment Table	0.012	0.010	0.002	0.008
	(0.020)	(0.020)	(0.020)	(0.025)
No. of obs.	83,029	83,029	82,391	18,868
Panel C: Health Index – ITT				
Born 1-3	0.029***	0.027**	0.028**	0.026*
	(0.011)	(0.011)	(0.011)	(0.014)
No. of obs.	87,336	87,336	86,207	19,757
Panel D: Health Index – LATE				
Treatment Table	0.052***	0.048**	0.048**	0.044*
	(0.019)	(0.019)	(0.019)	(0.024)
No. of obs.	87,336	87,336	86,207	19,757
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. In column two, we add year, month, and day of the week of birth fixed effects. The control variables added in columns (3) and (4) are maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

stronger long-run impacts for treated girls across both health and labor market outcomes (Appendix Table C.3). While we find economically small and insignificant estimates for education, we detect small positive impacts for girls on the average share of time in employment between ages 30-50 and a corresponding negative impact on the probability of being out of the labor force at age 50 (for the most part on disability pension or welfare benefits).[43] Around age 50 we observe all individuals in our outcome data, and this age is relevant as a non-trivial share of the population starts exiting the labor force (17 percent). Thus, our findings suggest

[43]Compliance with treatment assignment was similar across subgroups in our population. Thus, selective assignment of treatment should not drive these findings (Appendix Table B.4).

that the health index results bear relevance in the longer run also for economic outcomes, as these retirement decisions are predominantly due to health issues. In line with these findings for individual outcome measures, when we create separate indices for health, educational, and labor market outcomes in Appendix Table C.4 (the latter two constitute our combined SES index), we find that significant health and labor market impacts are driven by females (ITT impact of 0.022-0.028 standard deviations for the labor market index) even though we cannot formally reject equality across estimates for males and females.[44]

A factor ignored in our main analyses is the impact of extended nurse visits on mortality and emigration, both of which lead to individuals not being observed in our outcome data in either some or all years: First, non-merged nurse records may indicate early-life mortality. If extended nurse care increased the survival of weaker children, we may find attenuated impacts on the outcomes of survivors. Second, for the individuals merged to outcome data, we ignore that some individuals leave the data during the data period. To assess these issues, we perform two sets of analyses that (i) use a sample that includes non-matched nurse records and imputes individuals' mortality or emigration status or (ii) explicitly focus on long-run survivors by only considering individuals who are alive and in Denmark in 2017.

Table 3 examines mortality and emigration outcomes.[45] Columns (1) through (3), as well as (5) and (6), include all nurse records with imputed values (either a zero or a one) for non-merged individuals. Columns (4) and (7) use only merged individuals (our main analysis sample) and thus observed deaths or emigrations. Column (1) of Table 3 suggests that being born on the first three days of the month increases the probability of being observed in our outcome data by 0.4 percentage points at a baseline of 96.1 percent. In columns (2) through (7), we distinguish longer-run emigration and mortality outcomes. For ever having emigrated

[44]We do not find important fertility responses to the treatment and thus can rule out that this channel is important. 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 are not statistically significant. As exceptions, diagnoses for asthma and cancer stand out (Appendix Tables C.5 and C.6). In supplementary analyses, we have confirmed that our health index results are robust to the omission of asthma diagnoses and asthma-related hospital admissions, as well as alternative weighting schemes (Appendix Table C.7).

[45]Our measure for emigration is equal to one if an individual has ever left Denmark and was registered as an emigrant. It does not condition on never returning to Denmark. Mortality is an absorbing state.

Table 3 The Effect of Extended Nurse Care on Mortality and Ever Leaving Denmark (ITT).

	(1) Ever in data	(2) Ever Emi. imp. as 1	(3) Ever Emi. imp. as 0	(4) Ever Emi. obs	(5) Death imp. as 1	(6) Death imp. as 0	(7) Death 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: Each column shows estimates from a separate regression with the outcomes denoted as heading. The outcome in column (1) is an indicator for ever being observed in our merged outcome data, i.e., it is zero for non-merged nurse records. In columns (2) and (3), we examine the probability of ever emigrating from Denmark. We impute non-merged nurse records with either a one or a zero (i.e., we assume emigration/non-emigration for non-merged records). We only consider merged records and thus observed emigrations in column (4). Similarly, we analyze death by the latest outcome year (2020) in columns (5)-(7), with and without imputations of deaths for unmerged nurse records (imputed as "one" or "zero" in columns (5) and (6), respectively). As a result of the imputation exercises, the sample sizes in columns (2)-(3) and (5)-(6) are larger than for our main analyses (because we include individuals, who are not merged with the administrative data). Deaths and emigration (including temporary) are observed for all individuals in the 1968-2020 period. All regressions are based on our specification including year, month, and day of the week of birth fixed effects. MDV: Mean of the dependent variable in the control group. Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

we find no differences across treated and control children–also when imputing outcomes for individuals not observed in our outcome data. For mortality, our ITT estimate in column (7) suggests that children born on the first three days of the month are less likely to be dead by 2020, the final year of our mortality data (a seven percent change at the relevant mean in our merged sample). Imputing non-merged (and thus potentially dead) individuals as being dead increases our estimate for impacts on deaths by 2020.

In sum, our findings imply that children born on the first three days of the month likely had a higher probability of surviving childhood. Adult mortality impacts materialize at the final ages observed in our data, which is in line with mortality rates increasing from age 60. Three patterns in the data support this reasoning: First, mortality impacts are largest and only statistically significant for the oldest cohorts (1959-1961). Second, truncating our mortality data in the years 2016 or 2018, we find similar point estimates but they are imprecise, indicating power issues from lower death prevalence. Third, survival curves for the treated and the control group in Appendix Figure C.1 show emerging and significant survival differences at the oldest ages observed in our data.

As a final test of the robustness of our main results to selective attrition, we re-estimate our analyses for a sample of Danish residents in the year 2017, eliminating all individuals who either died or permanently emigrated. 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 C.8). In combination, our results imply that, if anything, we underestimate the positive health impacts of the extended nurse treatment.[46] "Weaker children" likely survived childhood due to the trial and thus the positive impacts on the good adult health index are consistent with health improvements outweighing this negative health selection. In the very long run, treated individuals see emerging mortality benefits that suggest that, if anything, longer-run follow-up will lead to larger estimates of the benefits of extended home visiting.

Heterogeneity across individuals and cohorts Examining heterogeneity of trial impacts can help assess two important questions: First, can universal programs play a role in alleviating health inequalities? Second, what is the role of the counterfactual environment for the importance of early life policies? The randomization that creates variation across days of the month of birth (rather than across cohorts) helps with both of these analyses.

We start with assessing heterogeneity of the impact of extended universal nurse visits across an initial condition easily observed by nurses: low birth weight status. To reduce the dimensionality in analyses of other relevant proxies for initial health and socioeconomic disadvantage, we also study heterogeneity across the summary measure of initial disadvantage.[47] Table 4 shows that, while we generally cannot reject equality of estimates across subgroups, we find larger health effects for children with initial disadvantages. For children with low birth weight, health effects are largest (and statistically different from the estimates in the non-low birth weight group at the 10 percent level when tested in a fully interacted model): Our instrumental variable estimates (0.197 standard deviations) document large im-

[46]In placebo analyses for individuals from three other major Danish towns, we do not find differential impacts of day of the month of birth on mortality in the 1977-2020 period or on the probability of ever leaving Denmark (Appendix Table C.9). As we only observe individuals from outside Copenhagen if they survive to 1977, the definitions of mortality and emigration are slightly different from our main analysis.

[47]Appendix Table C.10 presents additional separate heterogeneity analyses along dimensions of children's health, parental SES, and investment decisions in the first year.

pacts, which, evaluated at the mean of the outcome variable in the control group (-0.19), are economically meaningful. Our results suggest that extended nurse visiting compensated treated low birth weight children for the relative health disadvantage they face.[48] While point estimates for our SES index also suggest much larger impacts for individuals with low birth weight and initial disadvantages, those are imprecise.

Table 4 Individual-level Heterogeneity: The Effect of Extended Nurse Care on the SES and Health Indices across Subgroups, ITT and IV Estimates.

	Initial Disadvantage		Low BW Child	
	No	Yes	No	Yes
Panel A: SES Index (ITT and LATE)				
Born 1-3	-0.012	0.019	0.001	0.009
	(0.015)	(0.019)	(0.012)	(0.054)
P-value (Difference)	0.192		0.897	
Treatment Table	-0.020	0.036	0.003	0.015
	(0.024)	(0.036)	(0.021)	(0.094)
MDV	0.13	-0.18	0.01	-0.27
No. of obs.	49,152	33,239	78,745	4,075
Panel B: Health Index (ITT and LATE)				
Born 1-3	0.021	0.033*	0.021*	0.113**
	(0.014)	(0.018)	(0.011)	(0.054)
P-value (Difference)	0.605		0.093	
Treatment Table	0.034	0.064*	0.037*	0.197**
	(0.023)	(0.035)	(0.020)	(0.094)
MDV	0.05	-0.07	0.01	-0.19
No. of obs.	51,233	35,300	82,786	4,329

Notes: Each cell shows estimates from separate regressions estimated across subgroups of the population (denoted in the column headings). Initial disadvantage is a summary measure that is one if at least one of the following conditions holds: The mother was young at birth (below age 21), the child was born in a hospital, the father is missing in the administrative data, or the child was low birth weight. Regressions include year, month, and day of the week of birth fixed effects, but otherwise include no control variables. MDV is the mean of the dependent variable in the control group for the relevant subgroup. 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.

In a second heterogeneity analysis, we consider heterogeneity across trial cohorts. As shown in Figure 3, higher trial fidelity in initial years may help us understand the impact

[48] The average number of nurse visits for treated children does not differ by low birth weight status. However, we cannot rule out that other aspects, such as visit duration and intensity of advice, were adjusted.

Table 5 Cohort Heterogeneity: The Effect of Day of the Month of Birth on Outcome Indices by Trial Birth Cohorts (ITT), Sample of First Observed Child of Each Family in the Nurse Records.

Cohorts	1959-1961	1962-1964	1965-1967
Panel A: SES Index			
Born 1-3	0.031	0.005	-0.011
	(0.021)	(0.023)	(0.025)
No. of obs.	23,330	20,983	17,125
Panel B: Labor Market Index			
Born 1-3	0.044**	0.019	-0.012
	(0.021)	(0.023)	(0.024)
No. of obs.	23,654	21,163	17,165
Panel C: Education Index			
Born 1-3	0.007	0.004	-0.012
	(0.022)	(0.022)	(0.024)
No. of obs.	23,884	21,591	17,642
Panel D: Health Index			
Born 1-3	0.057**	0.058***	-0.025
	(0.023)	(0.022)	(0.022)
No. of obs.	24,406	21,971	17,869

Notes: Each column presents estimates from a separate regression of the respective outcome on an indicator for the child being born during the first three days of a month for birth cohorts in the nurse records defined in the table heading. We only include the first-observed child in the nurse records for each family. All regressions include year, month, and day of the week of birth fixed effects and the main set of control variables (maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data). Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

of the full treatment exposure vs partial exposure, with earlier trial cohorts experiencing higher compliance, especially for the third-year schedule. However, other changes over the trial period could play a role as well, most prominently a changing counterfactual policy environment. In particular, the expansion of universal childcare and other welfare services in the 1960s may play a role for the average impact of extended nurse care if all children benefit from other influential treatments that help compensate for initial disadvantage.

Table 5 documents stronger health impacts for the earliest trial cohorts that also extend to measures of SES for these children. We focus here on the first observed child for each family in the trial to make the composition of children across the cohort groups similar.[49] For initial trial cohorts, stronger impacts on health and labor market outcomes may point to the importance of a stringently implemented third year program or other important cohort differences. Cohort heterogeneity is consistent with the reasoning that important program elements such as health monitoring may be provided in different formats: Late trial cohorts were exposed to more policies with overlapping content, most prominently universal public childcare. As illustrated in Appendix Table A.1, childcare coverage in Copenhagen doubled during the trial periods, reaching around 50 percent for the 4-6 year-olds and effectively providing increasing shares of the control children with similar program components. Thus, this exposure likely impacted the counterfactual (children in our control group had potentially access to some treatments as well) and thus likely leads us to underestimate the importance of (any) toddler care provision for the final cohorts in the trial.

Spillovers in the Family Our main analyses have focused on the trial impacts on focal children. In the following, we factor in other family members to both understand the importance of toddler care and potential mechanisms. As nurses visited treated families during a period when they make decisions about aspects such as fertility, childcare take-up, and labor supply, a relevant mechanism for long-run impacts on children could be impacts on the fertility or labor supply of mothers. Appendix Table D.1 focuses 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). There are no economically or statistically significant impacts of the trial on fertility decisions or labor market outcomes. We conclude that it is unlikely that family adjustments at the margin of fertility or maternal labor supply drive our findings for long-run effects on the outcomes of focal children.

Moving to other family members, we study the role of sibling spillovers in the trial. Sibling spillovers, if present, may attenuate our findings in the main analyses because those ignore family ties. Two analyses suggest that this factor matters for our main results: First, when

[49] Keeping all children for the cohort analyses leads to very similar results. We discuss the potential importance of sibling spillovers in trial families in the next section.

we focus on a sample of only the first child observed in the trial period for each family (i.e., a sample that attempts to minimize family spillovers), we find larger trial impacts than in the main sample (Appendix Table D.2), albeit results for the SES index remain imprecise.

Second, Appendix Table D.3 zooms in on a sample of siblings. In these exploratory analyses, we regress siblings' outcomes on an indicator for the focal child being born on the first three days of the month, as well as focal child and sibling year-of-birth fixed effects.[50] We separately consider closely-spaced siblings (≤2 years), as they are particularly likely to be present in the family home during both their own first year visits and potential focal child extended nurse visits. For older siblings of treated focal children, Appendix Table D.3 shows, in general, very imprecise and often opposite signed impacts on the SES and health indices. Thus, we see no indication of spillovers to older siblings of focal children. For younger siblings, however, there appear to be large spillovers: Being the younger sibling of a treated focal child has a large positive impact on the younger sibling's SES index, especially if spacing is close. These siblings are likely to be in the family home when the extended treatment for the older focal child takes place. With respect to magnitude, the SES spillovers to younger siblings are precisely estimated and of similar magnitude as the impact of extended home visits on focal children themselves (comparing the estimate to Table 5 and the treatment effect in the oldest cohorts). Estimates for spillovers in health are imprecise but could be of similar size as our main results. These results suggest that our main results–that include siblings within families but only consider individuals' own treatment status–likely underestimate the treatment effects.

Exploring data from the nurse records in an even smaller sample of sibling-pairs with transcribed nurse records, we study potential channels for these spillovers. Younger, closely-spaced siblings of treated focal children appear more likely to have nurse registrations in their records (indicating that any spillovers may work through their own first-year nurse

[50] Recall that we focus on the first two children in multi-children families, and we require that we observe at least the focal (older or younger) child in each family to be observed with a nurse record. For further sample construction details, consult Section 3.2. All sibling analyses ignore the sibling's own treatment status for those individuals who are born in trial years themselves. Omitting siblings who are born in trial years and born on the first three days of the month does not impact our general conclusions.

treatment). Supporting the idea that close spacing makes nurse interactions with the entire family easier and thus benefits younger siblings, we also see suggestive evidence for closely-spaced younger siblings being more likely to be breastfed at one month than their counterparts with untreated older siblings. These exploratory analyses indicate that a more intensive first-year treatment for younger siblings may attenuate our main estimates. Our findings may also cautiously imply that more intensive first year contacts due to nurse visits for other children in the family (and resulting parental investments) are important for long-run SES outcomes.

Do the positive health impacts on treated children extend to their own offspring? In a final analysis, we study second-generation outcomes. Health at birth is a natural starting point, given extensive research documenting strong intergenerational ties and given its predictive power for later life outcomes. Intergenerational links in health at birth likely operate through both biological processes (defined in utero) and other channels. Those channels include socioeconomic status, partner choice, health status later in life, and lifestyle choices. Thus, we may expect that health investments in toddlers (that do not alter those children's own health at birth) may benefit their offspring especially through these channels.

Since we find the largest long-run health effects of extended nurse visits for first generation individuals with initial disadvantages, we hypothesize that any impacts on the second generation are most likely to be detected in the offspring of those first generation children. Focusing on children with low birth weight and whose offspring are themselves low birth weight is not feasible with our sample size and the share of treated individuals. Therefore, we focus on the sample of focal children who have an initial disadvantage more broadly defined. In our main second-generation analyses, we consider all children born to focal mothers or fathers to further increase sample size. However, we also present results for firstborn offspring.[51] For the second-generation children whose parents are both observed in the nurse data (when focal children have children together, 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, month and day of the week

[51]We define parity as the birth order of a given child within a family defined by the same mother.

of birth fixed effects, as well as focal child sex.[52]

Table 6 The Effect of Extended Nurse Care on Birth Outcomes in the Second Generation (ITT). Heterogeneity by Focal Child Initial Disadvantage

Second Gen.	Birth Weight (g)		Low BW		Preterm	
First Gen. Init. Disadv.	No	Yes	No	Yes	No	Yes
Panel A: All Second Gen. Children						
Born 1-3	-9.481	16.147*	0.001	-0.006*	0.005	-0.003
	(7.479)	(9.127)	(0.003)	(0.004)	(0.003)	(0.004)
MDV	3452.15	3383.56	0.05	0.07	0.06	0.07
No. of obs.	72,669	49,987	72,669	49,987	71,770	49,254
Panel B: Firstborn Second Gen. Children						
Born 1-3	-16.437	22.104*	-0.000	-0.012**	0.004	-0.004
	(10.326)	(12.660)	(0.004)	(0.005)	(0.005)	(0.006)
MDV	3367.75	3300.50	0.06	0.08	0.07	0.07
No. of obs.	35,917	24,342	35,917	24,342	35,352	23,845

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 firstborn children (Panel B). For second generation children with both parents (focal children) in our data, we keep only one spell for this child (the mother spell). All estimates are from our heterogeneity specification including fixed effects for focal child's year birth, month, 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. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table 6 shows our results for the impact of first generation treatment exposure on measures of second generation health at birth. We divide the sample of offspring by the initial disadvantage measure defined for the first generation. Analyses based on the full sample of focal children and their offspring result in imprecise estimates for all three birth outcomes. We find, however, suggestive evidence for the offspring of focal children with initial disadvantage benefitting from the trial exposure of their parent: We find a small increase in average birth weight and a decreased risk of being low birth weight. This effect is large: For second generation children with a parent with an initial disadvantage, the effect of this parent being born on the first three days of the month amounts to an 12 percent decrease

[52]Before we turn to second generation results, Appendix Table D.4 shows at most minimal impacts on focal children's fertility (the margin of having children at all, the number of children conditional on having a child, and the age at first birth). Effects are economically small and mostly not statistically significant.

of the risk of low birth weight (ITT effect). Effects for firstborn second generation children are qualitatively similar. Reassuringly, we find no effects on birth outcomes in a placebo sample of second generation children (born to disadvantaged parents) in the three largest towns outside Copenhagen (Appendix Table D.5).[53] Finally, our results are very similar but imprecise when we focus on smaller samples of the offspring of focal mothers or focal fathers separately (Appendix Table D.6). These findings suggest that the large health improvements we observe for focal children at the bottom of the health distribution may spill over to the health at birth of the second generation.[54]

Mechanisms: Childhood Outcomes To better understand the drivers of longer-run impacts of extended nurse visits, our final analyses examine childhood impacts of the extended nurse treatment. Figure 4 shows the intention-to-treat estimates and 95 percent confidence intervals for the impact of being born on the first three days of the month on a good health index at ages three and six, as well as a developmental milestone index at age three. For the milestones index, a negative estimate indicates earlier age at completion.

Figure 4 (and Appendix Figure C.2 for individual CPC measures) illustrates two main patterns: First, children born on the first three days of the month have better health during early childhood (with estimates for the childhood good health indices of 0.23 and 0.28 of a standard deviation, at ages three and six years). Digging into individual measures of health and health investments, we find no impact on full uptake of the suggested childhood vaccinations (our primary measure of uptake of preventive care), but a decrease in the likelihood to ever have consumed antibiotics during early childhood (measured for age groups 1-3 and 3-6

[53] We cannot perfectly match our measure of poor initial health in this sample as we do not observe first generation birth weights outside Copenhagen.

[54] As the second generation is still young, studying educational outcomes is challenging. Constraining our analysis to second generation children born prior to 1995, we consider a realized measure of adulthood educational attainment, completing more than compulsory education. On average, we find no second-generation impacts. We find some significant impacts on offspring educational attainment of being born on the first three days of the month, which are positive (negative) for children of focal individuals with (without) initial disadvantage (see Appendix Table D.7). However, these results are economically small (one percentage point at a control mean of 74-81 percent) and sensitive to our selection of second generation cohorts included. Thus, we cautiously conclude that there are no economically important education effects in the second generation.

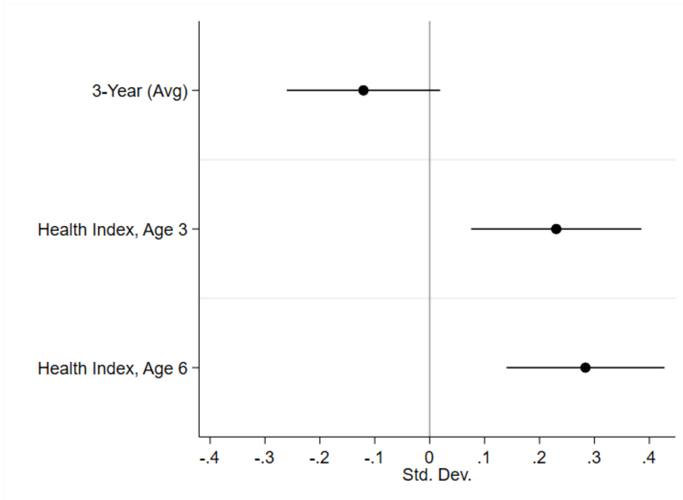


Fig. 4 Short-run Effects of Extended Nurse Care on Childhood Good Health Indices and a Developmental Milestones Index, ITT.

Notes: The figures present estimates from regressions of outcomes measured in the CPC (denoted on the Y-axis) on an indicator for being born on the first three days of the month (ITT), as well as 95 percent confidence intervals. The outcome "3-Year (Avg)" denotes a child's average score across all standardized developmental milestone scores measured at age three. Negative estimates for this outcome refer to an earlier age at milestone completion. See Section 3.3 and Appendix G for further details on outcome measures and sample construction.

years).[55] Given high prevalence rates of antibiotics consumption in our sample, this estimate likely reflects improvement in the general health of treated children (who needed less care in the health care sector for preventable infections) rather than improved access. Our estimates are large: Children born on 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 had been prescribed antibiotics at least once at age six). Scaling the ITT estimates by the implied first stage, we find that treated children see a 25 percentage points decrease in antibiotics use. This finding resonates with our longer-run results for asthma diagnoses, which have been related to childhood antibiotics

[55] Appendix Figure C.2 presents figures for vaccination completion and antibiotics usage, infectious disease prevalence and hospital admission events, standardized completion of developmental milestones across domains, and height. Appendix Tables C.11 through C.15 present the respective estimates (including and excluding a control for low birth weight status) and sample sizes across outcomes. We have also estimated all regressions separately for girls and boys. Point estimates are very similar but typically more precise for girls, with the exception of results for height during early childhood. While we find that height impacts in early years are driven by boys, we have reservations due to measurement issues in the height variable resulting in very small samples. Appendix Table C.11 confirms that the treatment does not predict whether a child is present in the CPC outcome data.

use (Patrick et al., 2020).[56]

Second, there is suggestive evidence for children born on the first three days of the month being younger at the completion of developmental milestones. However, estimates for the combined index and most subdomains remain imprecise. Thus, our results suggest that extended nurse visits were productive in promoting childhood health, while evidence along other dimensions is less clear. The developmental milestone measures may not be adequate to capture relevant dimensions of development, such as cognitive ability. Moreover, the dosage and content of the treatment may explain why we see limited impacts of extended nurse care.

5.4 Discussion of Effect Sizes, Relation to Existing Evidence, and Cost Effectiveness

How do our findings for adult effects of toddler preventive care (at around 0.05 standard deviations for our good health index) compare to those obtained for other childhood health policies with similar components? In general, those studies find much larger long-run health (and socioeconomic) impacts. This finding is likely due to the relevant counterfactual, the intensity, and target groups of the programs considered: The vast majority of (RCT and observational) empirical work considers high-intensity programs for disadvantaged families, a margin with scope for large impacts. Moreover, the program we study extended care, i.e., everybody had access to care in the first year.

Hoynes et al. (2016) find that early-life food stamp exposure in the 1960s to 1980s had significant and large impacts on adulthood metabolic health. Full in utero to age five exposure results in a 0.3 standard deviation reduction of metabolic syndrome in a high impact sample of poor families who were likely to have been very dependent on nutritional support. Scaled by the implied first stage, this effect translates into around 0.7 standard deviations for treated children (with large confidence intervals that include much smaller impacts). While food

[56] While additional estimates for hospital admissions and infectious disease are not significant at conventional levels in the full sample, similar to the long-run analyses, we find that some health benefits for girls during early childhood years are larger and significant at the 10 percent level.

stamps are a direct transfer and a more intensive nutrition policy than what we study, improved quality of toddler nutrition encouraged by nurses, especially among disadvantaged families, was at the core of the Danish program. 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 long-run returns, Bailey et al. (Forthcoming) 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. Full exposure during ages zero to five had the largest returns (an 11 percent decrease in mortality by the year 2012 for treated children at baseline mean of 4 percent). Our mortality estimates are of similar magnitude and thus arguably large (a decrease in mortality by 2020 of around 12.7 percent at a baseline mean of 8.3 percent).

Nurse visits were not only about nutritional advice but a composite treatment that encouraged other parental investments and uptake of additional care if necessary. Thus, our estimates are comparable to work focusing on the long-run effects of childhood exposure to health insurance (improving access to preventive care). Those studies primarily come from the US and focus on targeted rather than universal interventions. Wherry et al. (2018) find lower hospitalization rates at age 25 for children exposed to Medicaid around birth, 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 deviations for treated children, as well as in hospitalizations due to diabetes and obesity. These estimates are well in line with our smaller estimates for related health measures, especially in our sample of children with initial disadvantage.

Closely related to our work on toddler care, earlier work on the program roll out of first-year universal infant care in Denmark, Norway, and Sweden has documented health and mortality benefits of the introduction of these programs for the 1930s and 1940s cohorts (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).[57] Our work zooming in on follow-up care (set in the 1960s' 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 data requirements. Existing studies have examined targeted interventions using large-scale population data: prenatal exposure to Medicaid (East et al., 2023) and exposure to targeted early education programs/preschools (Rossin-Slater and Wüst, 2020; Barr and Gibbs, 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. (2023). They document large impacts on the risk of being low birth weight for first children 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 1.42 percentage points decrease for treated second-generation children in the probability of being low birth weight in our high-impact sample. Thus, our findings 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 of the second generation.

While much of the work referred to here documents average impacts of the studied interventions beyond health (including measures of educational attainment, labor market outcomes, neighborhood quality, or crime), our results for extended nurse visits are strongest for the health domain. As discussed, one explanation may be spillovers within the family that lead us to underestimate impacts on SES outcomes in our full sample. However, a larger health impact of the treatment (with consequences for labor market outcomes) may also

[57] Also closely related is 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).

relate to the following factors: First, our analyses on childhood data suggest that the health components of the program were particularly effective. Second, dosage and nurse qualifications are likely important: Even though nurses were supposed to also cover topics such as more general child development, our findings suggest that these components may not have been strong enough (dosage). Moreover, nurses had extensive training in the health domain but less experience with respect to toddler development in other domains. Third, children of the given cohorts were exposed to multiple programs in the expanding Danish welfare state, including universal childcare and later school doctor checkups, which may have served as partial substitutes and thus attenuate our findings, especially in the education domain. This finding is in line with earlier work on the interaction between targeted preschool and nurse home visit exposure in Denmark (Rossin-Slater and Wüst, 2020) and important when thinking about the generalizability of our results and their policy implications: Impactful elements of nurse treatment (presumably health monitoring, counseling, and advise) may be delivered in different programs, among them universal childcare with trained professionals.

Taking together all our results, we conclude that extended home visiting, a low cost intervention, was highly cost effective. Extrapolating from the figures on the default first-year program in Copenhagen and abstracting from fixed costs, we estimate the per-visit costs in the first-year program at DKK186/USD27 (2020).[58] Extrapolating to second- and third-year visits, we estimate per-child costs of only DKK931/USD133 (2020). Focusing only on the monetary value of deaths averted suggests that the trial benefits are larger than its costs by a factor of at least 10.[59] Thus, the long-run health benefits–especially among

[58]Nurses' yearly wages and allowances amounted to DKK47,000/DKK357,541 (1970/2020) (DNBH, 1970). They were responsible for around 160 children per year. We assume that those, on average, received 12 first-year home visits (as intended during the trial). The Statistical Yearbook for Copenhagen in 1963 reports a total of 56 nurses and a total of 8,879 infants under supervision (Copenhagen Statistical Office, 1963).

[59]Using our main estimate for the probability of death by 2020, we calculate that the trial prevented 55 deaths in the treatment group of 9,200 children (9,200*0.006). This figure suggests a cost per death averted of roughly USD22,000 (2020). We conservatively assume that this resulted in just one additional life year, valued at USD232,000 (as in Bailey et al. (2020)). We have attempted more sophisticated calculations of the mortality gains with respect to life expectancy impacts. However, we are constrained by sample size when estimating life expectancy changes across treated and control groups in our sample of roughly 92,000 individuals. As discussed, our results suggest that mortality differences across groups emerge at the oldest observed ages and thus we believe that our pragmatic focus on one additional life year is conservative but reasonable.

disadvantaged children–are likely to have greatly outweighed the very modest per-child costs of the program (which are admittedly lower bound estimates).

6 Conclusions

While causal evidence across settings documents that health shocks and targeted interventions matter for short- and long-run outcomes of disadvantaged infants, much less research has been conducted so far on the long-run and multi-generational impacts of universal policies for broader populations and policies that target older children. The design of these policies is likely to crucially matter for their returns. Understanding which elements of interventions work and how they interact with other programs is instrumental for policy makers in many contemporary settings with large-scale programs in place–much more so than evidence on the extensive margin of program introduction.

We contribute novel evidence on these topics by studying a large-scale government trial in Copenhagen that expanded the duration of a well-functioning infant home visiting program to also cover toddler years for a subset of children in the 1959-1967 cohorts. Our results document important long-run benefits for health in mid-adulthood and some positive SES impacts concentrated among girls and for labor market-related measures. Impacts are strongest for children of the earliest trial cohorts, who were likely unexposed to other influential treatments and likely exposed to the most stringently implemented extended program. As the cohorts enter age ranges with higher mortality risks, we also find positive impacts of toddler care on survival. Taken together, we document large returns for focal children of early-life preventive care, which we find plausibly works through improved health during childhood and better parental investments. Important for our understanding of home-based interventions such as nurse visiting, we show that some of the weaker longer-run impacts on SES measures in our main analyses may be due to family spillovers. These findings underline that accounting for benefits of home-based interventions for the entire family is crucial for understanding their returns.

We document that the health benefits of the program are particularly large for disadvantaged children, who see health benefits that compensate for their initial disadvantage and that may even spill over to the health of their offspring. This finding is notable, as extended care was not assigned to individuals based on measures of early disadvantage. Thus, universal policies may play an important role in mitigating early life inequalities. Even though targeting the program–and giving nurses a role to play in the assignment of families–could have further improved cost effectiveness, the second- and third-year visits came at very modest costs. Thus, the average impacts documented in our analyses would have justified implementation of the program as a universal offer, in particular for the initial trial cohorts.

Besides providing policy-relevant evidence on the long-run impacts of toddler-years universal preventive care, our study breaks new ground in making historical data on early-life circumstances available for analyses. In doing so, our paper points to exciting avenues for future work on topics such as intergenerational mobility, and for using transcribed historical data in combination with administrative data sources in Scandinavia and beyond.

References

Almond, Douglas, Janet Currie, and Valentina Duque (2018). “Childhood Circumstances and Adult Outcomes: Act II”. In: *Journal of Economic Literature* 56.4, pp. 1360–1446. DOI: 10.1257/jel.20171164.

Anders, John, Andrew Barr, and Alex Smith (2023). “The effect of early childhood education on adult criminality: Evidence from the 1960s through 1990s”. In: *American Economic Journal: Economic Policy* 15 (1), pp. 37–69.

Andersen, L. G., C. Holst, K. F. Michaelsen, J. L. Baker, and T. I. A. Sørensen (2012). “Weight and weight gain during early infancy predict childhood obesity: A case-cohort study”. In: *International Journal of Obesity* 36, pp. 1306–1311. DOI: https://doi.org/10.1038/ijo.2012.134.

Attanasio, Orazio, Ricardo Paes de Barros, Pedro Carneiro, David K. Evans, Lycia Lima, Pedro Olinto, and Norbert Schady (2022). *Public childcare, labor market outcomes of caregivers, and child development: experimental evidence from Brazil*. Tech. rep. National Bureau of Economic Research.

Attanasio, Orazio P. (Dec. 2015). “The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Policies”. In: *Journal of the European Economic Association* 13.6, pp. 949–997. ISSN: 1542-4766. DOI: 10.1111/jeea.12159.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey (2020). “How Well Do Automated Linking Methods Perform? Lessons from U.S. Historical Data”. In: *Journal of Economic Literature* 58, pp. 997–1044.

Bailey, Martha J., Hilary Hoynes, Maya Rossin-Slater, and Reed Walker (Forthcoming). “Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program”. In: *The Review of Economic Studies*. ISSN: 0034-6527. DOI: 10.1093/restud/rdad063.

Bailey, Martha J., Shuqiao Sun, and Brenden Timpe (2021). "Prep School for poor kids: The long-run impacts of Head Start on Human capital and economic self-sufficiency". In: *American Economic Review* 111.12, pp. 3963–4001.

Barker, David J. P. (1990). "The fetal and infant origins of adult disease". In: *BMJ: British Medical Journal* 301.6761, p. 1111.

Barr, Andrew, Jonathan Eggleston, and Alexander A. Smith (Apr. 2022). "Investing in Infants: The Lasting Effects of Cash Transfers to New Families". In: *The Quarterly Journal of Economics* 137 (4), pp. 2539–2583. ISSN: 0033-5533. DOI: `10.1093/qje/qjac023`.

Barr, Andrew and Chloe R. Gibbs (2022). "Breaking the Cycle? Intergenerational Effects of an Antipoverty Program in Early Childhood". In: *Journal of Political Economy* 130.12, pp. 3253–3285.

Barr, Andrew and Alexander A. Smith (2023). "Fighting crime in the cradle: The effects of early childhood access to nutritional assistance". In: *Journal of Human Resources* 58 (1), pp. 43–73. DOI: `https://doi.org/10.3368/jhr.58.3.0619-10276R2`.

Bartz, Christian, Hendrik Rätz, and Christoph Meinel (2021). "Handwriting Classification for the Analysis of Art-Historical Documents". In: *International Conference on Pattern Recognition* 314, pp. 562–576.

Belfield, Clive R., Milagros Nores, Steve Barnett, and Lawrence Schweinhart (2006). "The high/scope Perry Preschool Program cost–benefit analysis using data from the age-40 followup". In: *Journal of Human Resources* 41.1, pp. 162–190.

Besl, P. J. and Neil D. McKay (1992). "A method for registration of 3-D shapes". In: *Sensor Fusion IV: Control Paradigms and Data Structures*. Vol. 1611, pp. 586–607.

Bhalotra, Sonia, Martin Karlsson, and Therese Nilsson (2017). "Infant Health and Longevity: Evidence from A Historical Intervention in Sweden". In: *Journal of the European Economic Association* 15.5, pp. 1101–1157. ISSN: 1542-4766. DOI: `10.1093/jeea/jvw028`.

Bjerregaard, L. G., K. M. Rasmussen, K. F. Michaelsen, A. Skytthe, E. L. Mortensen, J. L. Baker, and T. I. A. Sørensen (2014). "Effects of body size and change in body size from infancy through childhood on body mass index in adulthood". In: *International Journal of Obesity* 38.10, pp. 1305–1311.

Bjerregaard, Lise G., Miriam Wüst, Torben SD Johansen, Thorkild IA Sørensen, Christian M Dahl, and Jennifer L Baker (Forthcoming). "Cohort Profile: the Copenhagen Infant Health Nurse Records (CIHNR) cohort." In: *International Journal Of Epidemiology*.

Bluche, Théodore, Jérôome Louradour, and Ronaldo Messina (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)*. Vol. 1. IEEE, pp. 1050–1055.

Bluche, Théodore, Hermann Ney, and Christopher Kermorvant (2014). "A Comparison of Sequence-Trained Deep Neural Networks and Recurrent Neural Networks Optical Modeling for Handwriting Recognition". In: *International Conference on Statistical Language and Speech Processing*. Ed. by Laurent Besacier, Adrian-Horia Dediu, and Carlos Martin-Vide. Cham.: Springer International Publishing, pp. 199–210.

Brock, Andy, Soham De, Samuel L. Smith, and Karen Simonyan (2021). "High-Performance Large-Scale Image Recognition Without Normalization". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, pp. 1059–1071.

Burbank, Allison J, Amika K Sood, Matthew J Kesic, David B Peden, and Michelle L Hernandez (2017). "Environmental determinants of allergy and asthma in early life". In: *Journal of Allergy and Clinical Immunology* 140.1, pp. 1–12.

Bütikofer, Aline, Katrine V. Løken, and Kjell G. Salvanes (2019). "Infant Health Care and Long-Term Outcomes". In: *The Review of Economics and Statistics* 101.2, pp. 341–354.

Clinchant, Stéphane, Hervé Déjean, Jean-Luc Meunier, Eva Maria Lang, and Florian Kleber (2018). "Comparing Machine Learning Approaches for Table Recognition in Historical Register Books". In: *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*. IEEE, pp. 133–138.

Copenhagen City Archives (1950-1970). *Archive Material from the 1950s and 1960s Copenhagen Nurse Home Visiting Program.*

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

Crimmins, Eileen M. and Caleb E. Finch (2006). "Infection, inflammation, height, and longevity". In: *Proceedings of the National Academy of Sciences* 103.2, pp. 498–503. ISSN: 0027-8424. DOI: 10.1073/pnas.0501470103.

Cubuk, Ekin D., Barret Zoph, Jonathon Shlens, and Quoc V. Le (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*, pp. 702–703.

Currie, Janet and Douglas Almond (2011). "Human capital development before age five". In: *Handbook of Labor Economics*. Ed. by Orley Ashenfelter and David Card. Vol. 4B. Amsterdam: Elsevier, North-Holland, pp. 1315–1486.

Currie, Janet and Enrico Moretti (2007). "Biology as destiny? Short -and long-run determinants of intergenerational transmission of birth weight". In: *Journal of Labor Economics* 25.2, pp. 231–264.

Dahl, Christian M., Torben S. D. Johansen, Emil N. Sørensen, Christian E. Westermann, and Simon F. Wittrock (2023a). "Applications of machine learning in tabular document digitisation". In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 56.1, pp. 34–48. DOI: 10.1080/01615440.2023.2164879.

Dahl, Christian M., Torben S. D. Johansen, Emil N. Sørensen, and Simon F. Wittrock (2023b). "HANA: A HAndwritten NAme Database for Offline Handwritten Text". In: *Explorations in Economic History* 87. Methodological Advances in the Extraction and Analysis of Historical Data, p. 101473. ISSN: 0014-4983. DOI: https://doi.org/10.1016/j.eeh.2022.101473.

Dahl, Christian M., Torben S. D. Johansen, Emil N. Sørensen, Christian E. Westermann, and Simon F. Wittrock (2022). "DARE: A large-scale handwritten DAte Recognition system". In: *arXiv preprint arXiv:2210.00503*.

Dahl, Christian M. and Christian E. Westermann (2023). "Tableparser: A deep learning based tool for segmentation of tabular documents". In: *Working Paper*.

Deng, Jia, Wei Dong, Richard Socher, L. Li, Kai Li, and Li Fei-Fei (2009). "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 248–255.

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]*. Tech. rep. The Danish National Board of Health.

Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby (2021). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". In: *ICLR*.

East, Chloe N., Sarah Miller, Marianne Page, and Laura R. Wherry (2023). "Multi-generational impacts of childhood access to the safety net: Early life exposure to Medicaid and the next generation's health". In: *American Economic Review* 113.1, pp. 98–135.

Ester, Martin, H. P. Kriegel, Jorg Sander, and Xiaowei Xu (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)*. AAAI Press, pp. 226–231.

Flensborg-Madsen, Trine and Erik L. Mortensen (2018). "Associations of early developmental milestones with adult intelligence". In: *Child development* 89.2, pp. 638–648.

Forsdahl, Anders (1979). "Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease?" In: *British Journal of Preventive and Social Medicine* 31.2, pp. 91–95.

Gao, Wei and Russ Tedrake (June 2019). “FilterReg: Robust and Efficient Probabilistic Point-Set Registration Using Gaussian Filter and Twist Parameterization”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11095–11104.

Geetha, R., T. Thilagam, and T. Padmavathy (2021). “Effective offline handwritten text recognition model based on a sequence-to-sequence approach with CNN–RNN networks”. In: *Neural Computing and Applications* 33.17, pp. 10923–10934.

Gern, James E., Robert F. Lemanske Jr., and William W. Busse (1999). “Early life origins of asthma”. In: *The Journal of Clinical Investigation* 104.7, pp. 837–843.

Goodfellow, Ian J., Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, and Vinay Shet (2013). “Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks”. In: *arXiv:1312.6082*.

Goodman-Bacon, Andrew (2018). “Public insurance and mortality: Evidence from Medicaid implementation”. In: *Journal of Political Economy* 126.1, pp. 216–262.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016). “Deep Residual Learning for Image Recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.

Heckman, James J. (2006). “Skill formation and the economics of investing in disadvantaged children”. In: *Science* 312.5782, pp. 1900–1902.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz (2010). “The rate of return to the HighScope Perry Preschool Program”. In: *Journal of Public Economics* 94.1-2, pp. 114–128.

Hjort, Jonas, Mikkel Sølvsten, and Miriam Wüst (2017). “Universal Investment in Infants and Long-Run Health: Evidence from Denmark’s 1937 Home Visiting Program”. In: *American Economic Journal: Applied Economics* 9.4, pp. 78–104. DOI: 10.1257/app.20150087.

Hoehn-Velasco, Lauren (2021). “The Long-term Impact of Preventative Public Health Programs”. In: *The Economic Journal* 131.634, pp. 797–826.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond (2016). "Long-run impacts of childhood access to the safety net". In: *American Economic Review* 106.4, pp. 903–34.

Huang, Gao, Yu Sun, Zhuang Liu, Daniel Sedra, and Kilian Q Weinberger (2016). "Deep networks with stochastic depth". In: *European Conference on Computer Vision*. Springer, pp. 646–661.

Kang, Lei, Pau Riba, Marçal Rusiñol, Alicia Fornés, and Mauricio Villegas (2022). "Pay Attention to What You Read: Non-recurrent Handwritten Text-Line Recognition". In: *Pattern Recognition* 129, p. 108766.

Kline, Patrick and Christopher R. Walters (July 2016). "Evaluating Public Programs with Close Substitutes: The Case of Head Start". In: *The Quarterly Journal of Economics* 131.4, pp. 1795–1848. ISSN: 0033-5533. DOI: 10.1093/qje/qjw027.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007). "Experimental analysis of neighborhood effects". In: *Econometrica* 75.1, pp. 83–119.

Kreiner, Claus Thustrup and Hans Henrik Sievertsen (2020). "Neonatal health of parents and cognitive development of children". In: *Journal of Health Economics* 69, p. 102247.

Lee, C. and Simon Osindero (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)*, pp. 2231–2239.

Merrick, J., T. W. Teasdale, and Y. Merrick (Dec. 1983). "School health screening of a birth cohort: A prospective longitudinal study". In: *International Journal of Rehabilitation Research* 6.4, pp. 461–468. ISSN: 0342-5282. DOI: 10.1097/00004356-198312000-00005.

Miller, Sarah and Laura R. Wherry (2019). "The long-term effects of early life Medicaid coverage". In: *Journal of Human Resources* 54.3, pp. 785–824.

Noghanibehambari, Hamid (2022). "Intergenerational health effects of Medicaid". In: *Economics & Human Biology* 45, p. 101114. ISSN: 1570-677X. DOI: https://doi.org/10.1016/j.ehb.2022.101114.

Olds, David L., Charles R. Henderson, Robert Chamberlin, and Robert Tatelbaum (1986). "Preventing Child Abuse and Neglect: A Randomized Trial of Nurse Home Visitation". In: *Pediatrics* 78.1, pp. 65–78.

Olds, David L., Charles R. Henderson, Robert Cole, John Eckenrode, Dennis Kitzmann Harriet Luckey, Lisa Pettitt, Kimberly Sidora, Pamela Morris, and Jane Powers (1998). "Long-term effects of nurse home visitation on childrens criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial". In: *JAMA* 280.14, pp. 1238–1244. DOI: 10.1001/jama.280.14.1238.

Olds, David L., Harriet Kitzman, Elizabeth Anson, Joyce A. Smith, Michael D. Knudtson, Ted Miller, Robert Cole, Christian Hopfer, and Gabriella Conti (2019). "Prenatal and infancy nurse home visiting effects on mothers: 18-year follow-up of a randomized trial". In: *Pediatrics* 144 (6), e20183889.

Patrick, David M. et al. (2020). "Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: Evidence from population-based and prospective cohort studies". In: *The Lancet Respiratory Medicine* 8.11, pp. 1094–1105.

Robling, Michael et al. (2016). "Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial". In: *The Lancet* 387.10014, pp. 146–155.

Rossin-Slater, Maya and Miriam Wüst (2020). "What Is the Added Value of Preschool for Poor Children? Long-Term and Intergenerational Impacts and Interactions with an Infant Health Intervention". In: *American Economic Journal: Applied Economics* 12.3, pp. 255–286.

Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei (2015). "ImageNet Large Scale Visual Recognition Challenge". In: *International Journal of Computer Vision* 115.3, pp. 211–252.

Schack-Nielsen, Lene, Thorkild I. A. Sørensen, Erik Lykke Mortensen, and Kim Fleischer Michaelsen (2010). "Late introduction of complementary feeding, rather than duration of breastfeeding, may protect against adult overweight”. In: *The American Journal of Clinical Nutrition* 91.3, pp. 619–627.

Shen, Zejiang, Ruochen Zhang, Melissa Dell, Benjamin Charles Germain Lee, Jacob Carlson, and Weining Li (2021). “LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis”. In: *International Conference on Document Analysis and Recognition*. Cham.: Springer, pp. 131–146.

Simonyan, Karen and Andrew Zisserman (2015). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. In: *International Conference on Learning Representations*, pp. 1–14.

Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov (2014). “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. In: *The Journal of Machine Learning Research* 15.1, pp. 1929–1958.

Statistics Denmark (2008). *60 år i tal. Danmark siden anden verdenskrig [60 years in numbers. Denmark since WWII]*. Tech. rep. Statistics Denmark.

Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna (June 2016). “Rethinking the Inception Architecture for Computer Vision”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826.

Tan, Mingxing and Quoc Le (Sept. 2019). “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”. In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, pp. 6105–6114.

— (2021). “EfficientNetV2: Smaller Models and Faster Training”. In: Proceedings of Machine Learning Research 139. Ed. by Marina Meila and Tong Zhang, pp. 10096–10106.

Touvron, Hugo, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Herve Jegou (2021a). “Training data-efficient image transformers & distillation through attention”. In: *Proceedings of the 38th International Conference on Machine Learning*. Ed.

by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, pp. 10347–10357.

Touvron, Hugo, Matthieu Cord, and Herve Jegou (2022). "DeiT III: Revenge of the ViT". In: *arXiv preprint arXiv:2204.07118*.

Touvron, Hugo, Matthieu Cord, Alexandre Sablayrolles, Gabriel Synnaeve, and Hervé Jégou (Oct. 2021b). "Going Deeper With Image Transformers". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 32–42.

Vallgårda, Signild (1996). "Hospitalization of deliveries: The change of place of birth in Denmark and Sweden from the late nineteenth century to 1970". In: *Medical history* 40.2, pp. 173–196.

van der Maaten, Laurens and Geoffrey Hinton (2008). "Visualizing data using t-SNE". In: *Journal of Machine Learning Research* 9.11, pp. 2579–2605.

Walker, Michael W et al. (Apr. 2023). *Intergenerational Child Mortality Impacts of Deworming: Experimental Evidence from Two Decades of the Kenya Life Panel Survey*. Working Paper 31162. National Bureau of Economic Research. DOI: `10.3386/w31162`.

Wherry, Laura R, Sarah Miller, Robert Kaestner, and Bruce D Meyer (2018). "Childhood medicaid coverage and later-life health care utilization". In: *Review of Economics and Statistics* 100.2, pp. 287–302.

Wightman, Ross (2019). *PyTorch Image Models*. `https://github.com/rwightman/pytorch-image-models`. DOI: `10.5281/zenodo.4414861`.

Wightman, Ross, Hugo Touvron, and Hervé Jégou (2021). "ResNet strikes back: An improved training procedure in timm". In: *arXiv preprint arXiv:2110.00476*.

Wüst, Miriam (2012). "Early interventions and infant health: Evidence from the Danish home visiting program". In: *Labour Economics* 19.4, pp. 484–495.

Wüst, Miriam (May 2022). "Universal Early-Life Health Policies in the Nordic Countries". In: *Journal of Economic Perspectives* 36.2, pp. 175–98. DOI: `10.1257/jep.36.2.175`.

You, Yang, Jing Li, Sashank Reddi, Jonathan Hseu, Sanjiv Kumar, Srinadh Bhojanapalli, Xiaodan Song, James Demmel, Kurt Keutzer, and Cho-Jui Hsieh (2019). "Large Batch

Optimization for Deep Learning: Training BERT in 76 minutes”. In: *arXiv preprint arXiv:1904.00962*.

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

Zhong, Zhun, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang (2020). “Random Erasing Data Augmentation”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 07, pp. 13001–13008.

Online Appendices

A Data and Descriptive Statistics

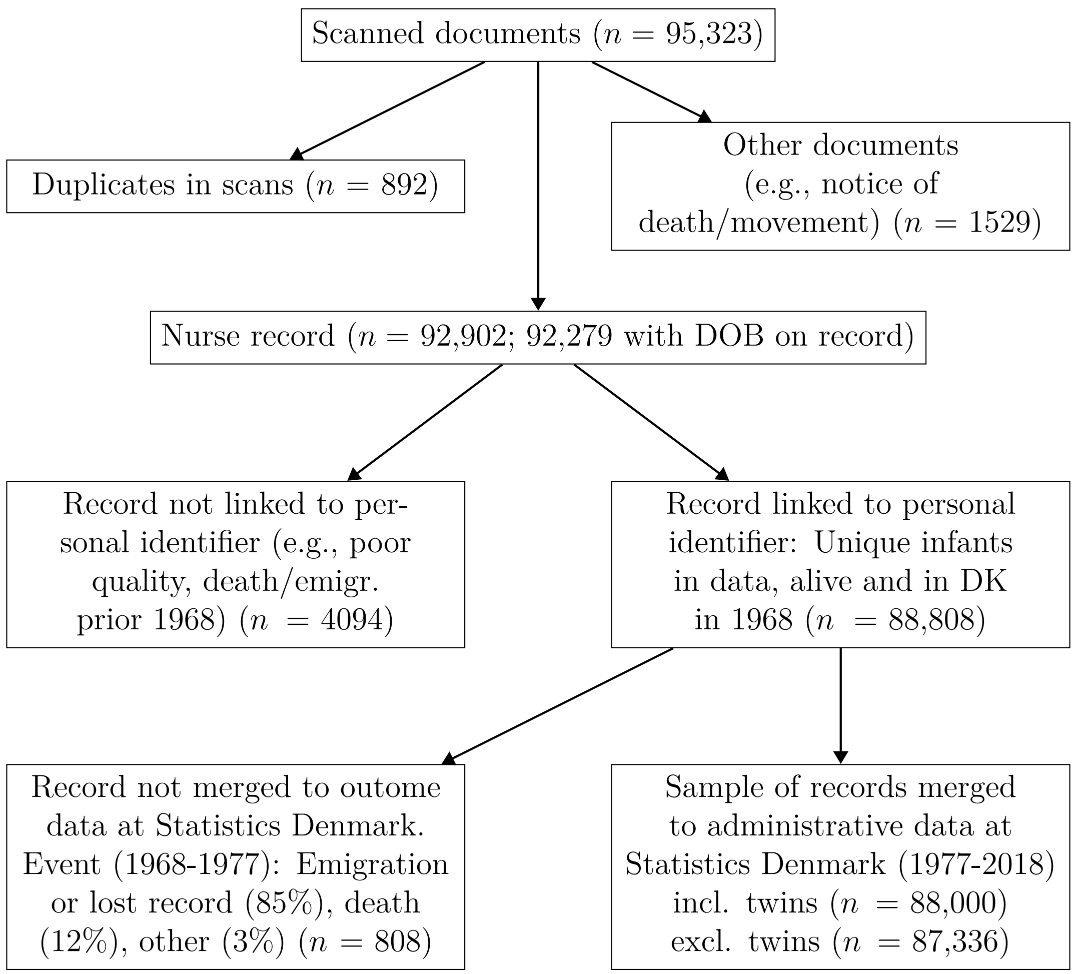
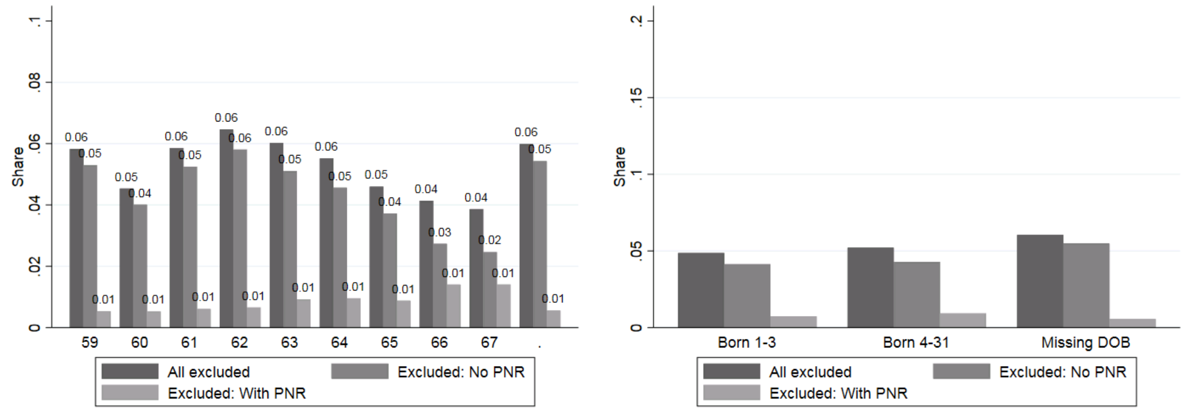


Fig. A.1 Flowchart: Linkage and Merge of Scanned Records to Administrative Data.
Notes: Numbers in parentheses indicate sample sizes. DOB: date of birth.



(a) Excluded by year

(b) Excluded by day of the month (groups)

Fig. A.2 Share of Excluded Nurse Records (with and without a unique personal identifier), by Birth Cohort and by Day of the Month of Birth, 1959-1967.

Notes: PNR is the personal identifier assigned to all Danish residents from 1968 on. DOB: date of birth.

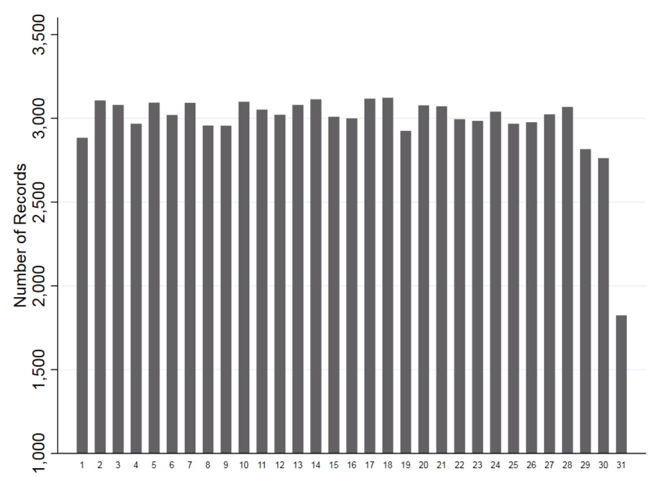


Fig. A.3 Number of Records across Days of the Month, 1959-1967.

Notes: This figure includes all nurse records irrespective of their match status with outcome data. It shows the frequency of nurse records across the days of month of birth.

Table A.1 Number of Children in Copenhagen (Aged 0-6) Enrolled in Public Childcare and Total Number of Children in Copenhagen, Selected Years.

	1959	1963	1966	1968	1970
Children in Pub. Childcare, Age 0-3 (Vuggestuer)	1,581	1,626	1,575	2,822	3,173
Children in Pub. Childcare, Age 4-6 (Børnehaver)	7,980	8,123	8,212	11,836	12,221
Total No. of Children in Pub. Childcare	9,561	9,749	9,787	14,658	15,394
No. of Resident Children, Age 0-6	61,947	59,625	57,302	53,872	50,441
Estimated Coverage (%)	15	16	17	27	31

Notes: The table is based on aggregate statistics reported in the Statistical Yearbooks for Copenhagen (Copenhagen Statistical Office, various years). Population counts for 1963 and 1968 are imputed as simple averages of the prior and later years reported in the table (due to population counts being based on censuses in the years 1960, 1966, and 1970.) The estimated coverage rate divides the number of enrolled children by the population count. The yearbook for the year 1972 reports coverage rates of 19.5 and 49.1 for the 0-3 and 3-6 age groups, respectively. Earlier yearbooks do not report coverage rates.

Table A.2 Content of the First Year Home Visiting Program in Copenhagen.

		Age of Child		
Topic	Example Items	2 weeks	1, 2, 3, 4, 6, 9 months	12 months
Health	Weight	✓	✓	✓
	General health assessment, illnesses in first year, hospital admissions			✓
Health care take-up	GP care uptake, vaccination status, ever hospitalized			✓
Nutrition	Feeding mode, number of meals	✓	✓	✓
	Duration of breastfeeding			✓
Infant development	Smiles, lifts head, babbles, sits alone		✓	✓
	Height, walks, number of teeth			✓
	Child has own bed, hygiene	✓	✓	✓
Family	Socioeconomic status, childcare attendance/mode of care		✓	✓
Mother	Employment status, mother mental and physical health		✓	✓

Notes: The table shows topics covered in the first-year visits and example items for nurse registrations in the children'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 City Archives, various years). For each age-specific topic, nurse registrations were made at one of those visits.

Table A.3 Content of the Home Visiting Program in Copenhagen, Second and Third Year.

Topic	Example Items	Age of Child 15 months	18 months	24 months	30 months	36 months
Health and accidents	Diagnosed illnesses, dyspepsia, lung illnesses, otitis, other, accidents	✓	✓	✓	✓	✓
Health care take-up	GP consultations, uptake of dental care, vaccinations, hospital admissions	✓	✓	✓	✓	✓
Nutrition and teeth	Vitamins, candy, healthy diet, teeth brushing	✓		✓		✓
Child development	Emotional problems, habit formation, hygiene, child has own bed		✓	✓	✓	✓
	Language, gross-motor					✓
Parenting style, childcare	Strictness, parent-child relationship, childcare attendance					✓

Notes: The table shows topics covered in the second and third year nurse program and example items for nurse registrations in the children's records. The extended program offered five visits during the child's second and third year of life.

Table A.4 Descriptive Statistics (Means, Std.Dev.), CPC Sample.

	Full Sample	Born 4-31	Born 1-3	N	p-value
A. Background Characteristics, CPC Analysis Sample at Three Years					
Female	0.49 (0.50)	0.49 (0.50)	0.43 (0.50)	2,704	0.06
Social status (0 = low, 1 = middle, 2 = high)	1.15 (0.82)	1.15 (0.82)	1.13 (0.83)	2,314	0.69
Smoking in last trimester of pregnancy	0.53 (0.50)	0.54 (0.50)	0.52 (0.50)	2,655	0.55
Preterm birth	0.17 (0.37)	0.17 (0.38)	0.14 (0.35)	2,223	0.24
Birth order	1.76 (1.12)	1.75 (1.12)	1.83 (1.10)	2,704	0.26
Birth weight (g)	3214.09 (576.46)	3209.45 (583.62)	3257.69 (503.08)	2,704	0.20
Low BW	0.09 (0.29)	0.10 (0.29)	0.06 (0.23)	2,704	0.04
B. Outcome Measures, CPC Analysis Samples at Three and Six Years					
Child observed in 3yr examination	0.62 (0.49)	0.62 (0.48)	0.59 (0.49)	4,369	0.16
Height (cm), age 3	95.89 (4.06)	95.81 (4.03)	96.59 (4.22)	1,206	0.04
Height (cm), age 6	118.55 (5.48)	118.43 (5.43)	119.56 (5.88)	1,090	0.04
Height (cm), age 7	122.92 (5.31)	122.86 (5.33)	123.39 (5.10)	1,773	0.22
Height (cm), age 10	138.15 (6.35)	138.13 (6.31)	138.27 (6.72)	2,400	0.74
Height (cm), age 13	155.43 (7.91)	155.48 (7.91)	154.95 (7.98)	2,348	0.33
Numbers of vaccinations, age 3	2.41 (0.90)	2.40 (0.90)	2.42 (0.95)	2,402	0.73
Numbers of vaccinations, age 6	2.05 (1.03)	2.05 (1.03)	2.09 (1.02)	2,716	0.58
Fully vaccinated, age 3	0.60 (0.49)	0.60 (0.49)	0.65 (0.48)	2,402	0.16
Fully vaccinated, age 6	0.44 (0.50)	0.44 (0.50)	0.46 (0.50)	2,716	0.50
Ever consumed antibiotics, age 1-3	0.48 (0.50)	0.48 (0.50)	0.39 (0.49)	2,273	0.01
Ever consumed antibiotics, age 3-6	0.47 (0.50)	0.48 (0.50)	0.34 (0.47)	1,888	0.00
Any hospital admission, age 1-3	0.27 (0.44)	0.27 (0.44)	0.25 (0.43)	2,402	0.52
Any hospital admission, age 3-6	0.32 (0.47)	0.33 (0.47)	0.29 (0.46)	2,716	0.24
Otitis, bronchitis or pneumonia, age 1-3	0.30 (0.46)	0.30 (0.46)	0.26 (0.44)	2,704	0.13
Otitis, bronchitis or pneumonia, age 3-6	0.31 (0.46)	0.31 (0.46)	0.28 (0.45)	2,743	0.22
Mean milestone z-score: Language	-0.00 (1.00)	0.01 (1.00)	-0.05 (1.01)	2,001	0.48
Mean milestone z-score: Motor	0.00 (1.00)	0.01 (1.00)	-0.09 (0.98)	2,180	0.15
Mean milestone z-score: Eating	-0.00 (1.00)	0.01 (1.00)	-0.14 (1.00)	1,946	0.05
Mean milestone z-score: Dressing	0.00 (1.00)	-0.00 (1.00)	0.02 (1.03)	1,698	0.76
Mean milestone z-score: Soc. Interaction	-0.00 (1.00)	0.02 (0.99)	-0.16 (1.07)	2,021	0.02
Mean milestone z-score: Toilet	0.00 (1.00)	0.00 (1.00)	-0.02 (1.02)	2,134	0.80
Mean milestone z-score: Combined	-0.00 (1.00)	0.01 (1.00)	-0.11 (0.99)	2,256	0.09

Notes: Summary statistics for the CPC sample with a valid registration for the relevant outcome measure. For background characteristics, we report descriptives for children who participated at least in the three year examination. The height measures beyond age six are from school doctor measurements. The p-value in the final column is from a t-test of equality of means.

Table A.5 Descriptive Statistics (means, std.dev.), Background Characteristics for Children Born in Copenhagen and Children Born outside Copenhagen, 1959-1967.

	All Births Sample	Outside CPH	CPH	No. of Obs.	p-value
Female	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	694,094	0.27
Hospital Birth	0.07 (0.26)	0.03 (0.17)	0.32 (0.47)	694,094	0.00
Family Size	1.89 (0.86)	1.91 (0.87)	1.77 (0.78)	688,983	0.00
Mother's Age at Birth	26.46 (5.70)	26.59 (5.70)	25.70 (5.67)	688,983	0.00
Firstborn	0.45 (0.50)	0.43 (0.50)	0.53 (0.50)	688,983	0.00
Father Missing	0.03 (0.16)	0.02 (0.15)	0.05 (0.21)	692,299	0.00

Notes: The table displays summary statistics for all births in the given sample (defined by place of birth registration) and observed in the administrative data at Statistics Denmark in 1980 or later. Importantly, births do not equal residents. The p-value in the final column is from a t-test of equality of means.

B Robustness of the First Stage

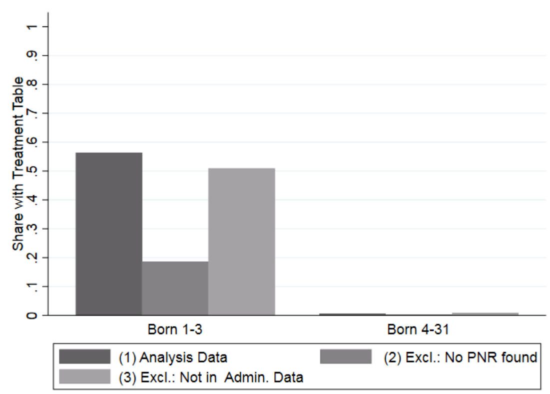


Fig. B.1 Share of Records with a Treatment Table by Day of the Month of Birth. Matched and Unmatched Nurse Records.

Notes: PNR is the personal identifier assigned to all Danish residents from 1968 on.

Table B.1 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 geographical units in our sample. Areas (*provsti*) are defined by birth registration codes, our best proxy for residence. *Other* summarizes individuals with unknown registration codes, religious minorities, and with birth registrations outside Copenhagen. Together with hospital birth codes, those do not belong to a geographically defined unit in Copenhagen.

Table B.2 First Stage: Presence of a Treatment Table and Completeness of Registrations. Main Analysis Sample.

	(1)	(2)	(3)	(4)
Panel A: Treatment Table Present				
Born 1-3	0.568***	0.568***	0.574***	0.574***
	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	87,336	87,336	86,207	19,757
F-value	11,330	440	356	360
Panel B: Any Cell of Treatment Table Filled				
Born 1-3	0.502***	0.502***	0.508***	0.509***
	(0.006)	(0.006)	(0.006)	(0.006)
No. of obs.	86,632	86,632	85,520	19,072
F-value	7,987	311	249	251
Panel C: Share of Cells of Treatment Table Filled				
Born 1-3	0.394***	0.394***	0.399***	0.399***
	(0.005)	(0.005)	(0.005)	(0.005)
No. of obs.	86,632	86,632	85,520	19,072
F-value	6,412	249	198	200
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. The control variables in columns (3) and (4) are maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table B.3 First Stage: Presence of a Treatment Table Across Samples.

	(1)	(2)	(3)	(4)
Panel A: All Nurse Records, including Non-Matched Records				
Born 1-3	0.552***	0.565***	0.645***	0.646***
	(0.005)	(0.005)	(0.008)	(0.008)
No. of obs.	92,279	88,009	34,555	8,035
F-value	11,186	441	193	196
Panel B: Non-Empty Nurse Records at One Year				
Born 1-3	0.577***	0.577***	0.647***	0.647***
	(0.005)	(0.005)	(0.007)	(0.007)
No. of obs.	86,284	86,284	47,054	10,816
F-value	11,449	444	232	238
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

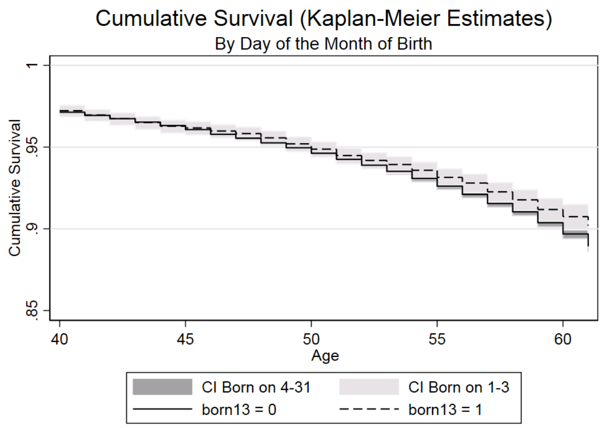
Notes: Each cell shows estimates from a separate regression. Panel (A) uses the full sample of transcribed nurse records, including non-matched records. Panel (B) uses our main sample of matched nurse records and conditions further on the nurse record being non-empty for registrations at the 12 months visit. Cpntrols in Panel (B) are the same as in the main analyses. Controls in Panel (A) include a set of variables from the transcribed nurse records (indicators for a birth prior to due date, for low birth weight, for nurse evaluation of a good economic status of the family, for good maternal mental and physical health, for observed harmony in the family, for an orderly home, and for exclusive breastfeeding, all recorded at child age one month). Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table B.4 First Stage: Presence of a Treatment Table for Subgroups of the Population, Main Analysis Sample.

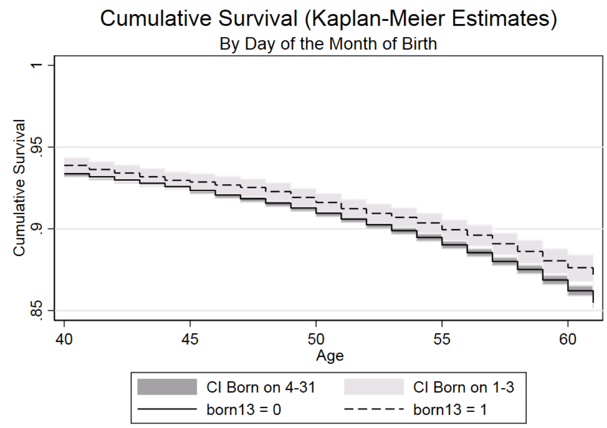
	(1) Initial Disadvantage		(2) Low BW		(3) Firstborn		(4) Mother's age $\geq$ Median (24 yrs)		(5) Hosp. birth		(6) Miss. father		(7) Mother $\geq$ comp. edu.		(8) Breastfed	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Born 1-3	0.614***	0.515***	0.574***	0.572***	0.610***	0.534***	0.531***	0.601***	0.593***	0.475***	0.572***	0.496***	0.583***	0.555***	0.593***	0.628***
	(0.007)	(0.009)	(0.006)	(0.024)	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.012)	(0.005)	(0.023)	(0.007)	(0.008)	(0.009)	(0.007)
Ratio to full pop.	1.08	0.91	1.01	1.01	1.07	0.94	0.93	1.06	1.04	0.84	1.01	0.87	1.03	0.98	1.05	1.11
No. of obs.	51,233	35,300	82,786	4,329	38,868	47,886	41,022	45,732	68,237	19,099	82,591	4,745	44,342	39,061	33,099	44,677
P-value (Difference)	0.000		0.949		0.000		0.000		0.000		0.001		0.010		0.002	

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 year, month, and day of week of birth fixed effects, but exclude other control variables. Breastfed equals one if the child was registered by a nurse as being breasted at age one month. The *ratio to full pop.* 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.

C Robustness of Main Results and Mechanisms



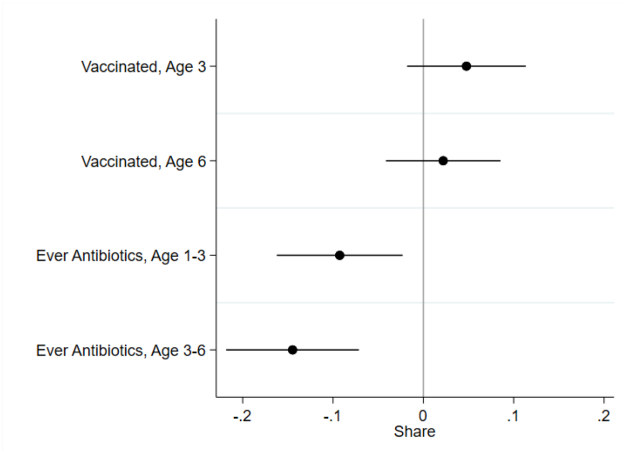
(a) Deaths Observed in Outcome Data



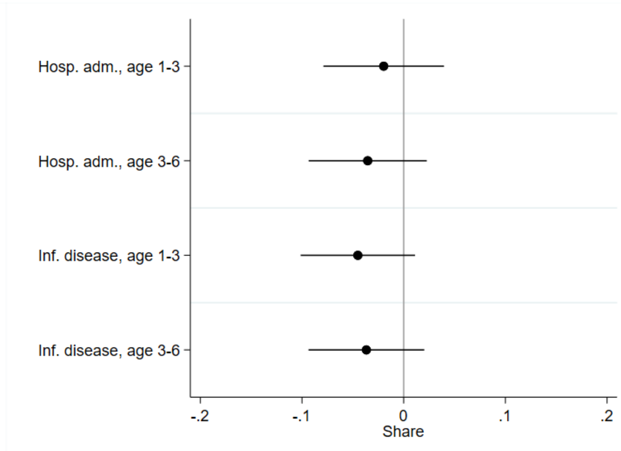
(b) Imputed Deaths for Non-Linked Nurse Records

Fig. C.1 Survival Estimates (Kaplan-Meier) by Day of the Month of Birth.

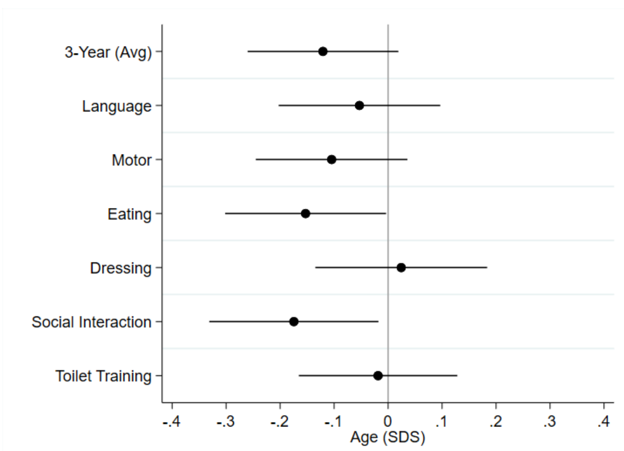
Notes: The figures show survival curves and 95 percent confidence intervals for our main sample with observed deaths in the period 1970-2020 (a) and including non-linked and non-merged records with imputed deaths by 1970 (b). By the end of our data period, individuals are between 53 and 61 years. Across ages, we focus on the day of the month of birth comparison.



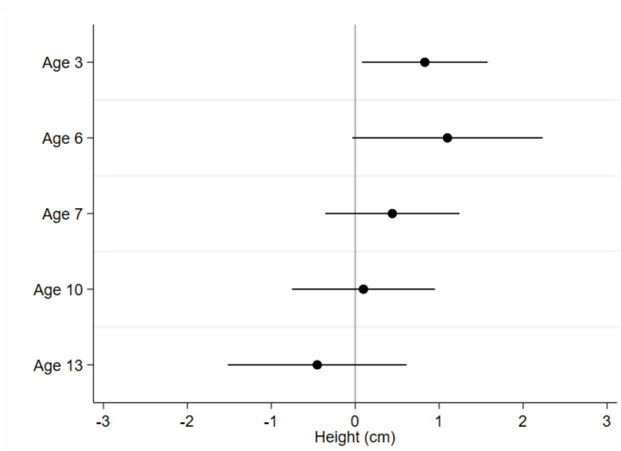
(a) Vacc. Status and Antibiotics Use



(b) Hosp. Adm. and Inf. Disease



(c) Dev. Milestones



(d) Height

Fig. C.2 Short-run Effects of Extended Nurse Care on Health, Health Care Use, and Developmental Milestones across Domains, ITT.

Notes: The figures present estimates from regressions of outcomes (denoted on the Y-axis) on an indicator for being born on the first three days of the month (ITT), and 95 percent confidence intervals. Regressions for height include a linear control for age at measurement. In Panel (c), "3-Year (Avg)" denotes a child's average score across all standardized developmental milestone scores. See Section 3.3 and Appendix G for further details.

Table C.1 Placebo: The Effect of Day of the Month of Birth on the SES and Health Indices in the Danish Towns Sample (ITT), 1959-1967.

	(1)	(2)	(3)	(4)
Panel A: SES Index				
Born 1-3	0.011	0.011	0.007	0.007
	(0.013)	(0.013)	(0.013)	(0.016)
No. of obs.	64,197	64,197	63,756	14,724
Panel B: Health Index				
Born 1-3	0.007	0.007	0.005	0.015
	(0.013)	(0.013)	(0.013)	(0.016)
No. of obs.	66,727	66,727	65,997	15,220
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. The controls added in columns (3) and (4) include maternal age at birth and indicators for child sex, the child being born in a hospital, and the father not being observed in the administrative data. We do not control birth weight, as we do not observe nurse registrations in the towns sample (Aarhus, Odense, and Aalborg). We add town fixed effects additional to year, month, and day of the week of birth fixed effects. Robust standard errors are in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.2 Power Calculations for First Generation Long-Run Outcomes.

Variable	Req. Effect Size	No. of Obs.
Yrs. of educ.	0.082	85,330
Above compulsory educ.	0.014	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. Earnings are in 2015 DKK. The diagnoses record whether an individual has ever been diagnosed with the illness.

Table C.3 The Effects of Day of the Month of Birth on Education, Income, and Labor Market Status, Full Sample and By Gender, ITT.

	Full Sample	Females	Males
Panel A: Years of Education			
Born 1-3	-0.028	-0.000	-0.058
	(0.029)	(0.039)	(0.042)
No. of obs.	84,652	41,609	43,043
Panel B: Above Compulsory Education			
Born 1-3	0.000	0.009	-0.009
	(0.005)	(0.007)	(0.007)
No. of obs.	84,652	41,609	43,043
Panel C: Higher Education			
Born 1-3	-0.005	-0.004	-0.006
	(0.005)	(0.008)	(0.007)
No. of obs.	84,652	41,609	43,043
Panel D: University Education			
Born 1-3	-0.004	-0.004	-0.003
	(0.003)	(0.005)	(0.005)
No. of obs.	84,652	41,609	43,043
Panel E: Earnings at 25 (2015 DKK)			
Born 1-3	313.550	-246.199	887.582
	(1517.278)	(1962.614)	(2305.906)
No. of obs.	85,037	41,629	43,408
Panel F: Average Earnings 30-50 (2015 DKK)			
Born 1-3	1264.160	3199.164	-722.962
	(2044.914)	(2391.656)	(3299.936)
No. of obs.	84,329	41,316	43,013
Panel G: Share in Employment Age 30-50			
Born 1-3	0.006*	0.014***	-0.003
	(0.003)	(0.005)	(0.005)
No. of obs.	83,436	40,880	42,556
Panel H: Out of the Labor Force, Age 50			
Born 1-3	-0.006	-0.017***	0.005
	(0.004)	(0.006)	(0.006)
No. of obs.	78,000	38,893	39,107

Notes: Each cell shows estimates from a separate regression. We apply our main specification with year, month, and day of the week of birth fixed effects as well as controls (maternal age at birth and indicators for child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data). Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table C.4 The Effect of Day of the Month of Birth on Health, Education, and Labor Market Indices, Samples of Females and Males, ITT.

	(1)	(2)	(3)	(4)
Panel A: Health Index, Females				
Born 1-3	0.035**	0.033**	0.032**	0.035*
	(0.015)	(0.015)	(0.015)	(0.020)
No. of obs.	42,690	42,690	42,138	9,663
Panel B: Health Index, Males				
Born 1-3	0.026*	0.025	0.024	0.017
	(0.016)	(0.016)	(0.016)	(0.020)
No. of obs.	44,646	44,646	44,069	10,094
Panel C: Education Index, Females				
Born 1-3	0.012	0.010	0.005	0.011
	(0.016)	(0.016)	(0.016)	(0.020)
No. of obs.	41,954	41,954	41,609	9,556
Panel D: Education Index, Males				
Born 1-3	-0.018	-0.018	-0.022	-0.009
	(0.016)	(0.016)	(0.016)	(0.020)
No. of obs.	43,376	43,376	43,043	9,829
Panel E: Labor Market Index, Females				
Born 1-3	0.027*	0.028*	0.025*	0.022
	(0.014)	(0.014)	(0.014)	(0.018)
No. of obs.	41,086	41,086	40,756	9,330
Panel F: Labor Market Index, Males				
Born 1-3	-0.000	-0.001	-0.004	0.011
	(0.017)	(0.017)	(0.017)	(0.022)
No. of obs.	42,765	42,765	42,446	9,725
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. The control variables added in columns (3) and (4) are maternal age at birth and indicators for child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.5 The Effect of Day of the Month of Birth on Diagnoses Outcomes, Full Sample and By Gender, ITT.

	Full Sample	Females	Males
Panel A: Diabetes			
Born 1-3	-0.002	-0.005*	-0.000
	(0.002)	(0.003)	(0.003)
No. of obs.	86,207	42,138	44,069
Panel B: Cardiovascular Disease			
Born 1-3	-0.005	-0.010	-0.001
	(0.005)	(0.007)	(0.007)
No. of obs.	86,207	42,138	44,069
Panel C: Heart Disease			
Born 1-3	-0.002	-0.004	0.001
	(0.003)	(0.003)	(0.004)
No. of obs.	86,207	42,138	44,069
Panel D: Asthma			
Born 1-3	-0.005**	-0.005	-0.005*
	(0.002)	(0.004)	(0.003)
No. of obs.	86,207	42,138	44,069
Panel E: Cancer			
Born 1-3	-0.006**	-0.007	-0.006
	(0.003)	(0.005)	(0.004)
No. of obs.	86,207	42,138	44,069
Panel F: Mental Health Issues			
Born 1-3	-0.001	0.004	-0.006
	(0.004)	(0.006)	(0.005)
No. of obs.	86,207	42,138	44,069
Panel G: Infection			
Born 1-3	-0.000	0.000	-0.001
	(0.005)	(0.007)	(0.006)
No. of obs.	86,207	42,138	44,069

Notes: Each cell shows estimates from a separate regression. All diagnosis outcomes measure the probability of ever being observed with the given diagnosis in our outcome data. We apply our main specification with year, month, and day of the week of birth fixed effects, as well as controls (maternal age at birth and indicators for child sex (we omit child sex as a control in the last two columns), child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data). Robust standard errors in parentheses.* p<0.1 ** p<0.05 *** p<0.01.

Table C.6 The Effect of Day of the Month of Birth on Hospitalization Outcomes, Full Sample and By Gender, ITT.

	Full Sample	Females	Males
Panel A: Hospital Nights, Age 30-39			
Born 1-3	-0.205	-0.190	-0.226
	(0.206)	(0.315)	(0.266)
No. of obs.	86,207	42,138	44,069
Panel B: Hospital Nights, Age 40-49			
Born 1-3	-0.307	-0.574*	-0.052
	(0.228)	(0.336)	(0.308)
No. of obs.	86,207	42,138	44,069
Panel C: Hospital Nights, Age 50-59			
Born 1-3	-0.061	-0.256	0.123
	(0.194)	(0.271)	(0.276)
No. of obs.	86,207	42,138	44,069
Panel D: Ever Hospitalized, Age 30-59			
Born 1-3	-0.007	-0.006	-0.008
	(0.005)	(0.006)	(0.008)
No. of obs.	86,207	42,138	44,069

Notes: Each cell shows estimates from a separate regression. We apply our main specification with year, month, and day of the week of birth fixed effects, as well as controls (maternal age at birth and indicators for child sex (we omit child sex as a control in the last two columns), child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data). Robust standard errors in parentheses.* $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table C.7 Robustness: Alternative Good Health Indices, ITT.

	Health Index Main	Health Index Equ. Weight	Health Index No Asthma
Born 1-3	0.028**	0.026**	0.023**
	(0.011)	(0.011)	(0.011)
No. of obs.	86,207	86,207	86,207

Notes: Each cell shows estimates from a separate regression for the outcome denoted in the column head (our main good health index, an equally weighted health index, and a health index excluding asthma hospitalizations and diagnoses). For the equally weighted health index, we weigh diagnoses and hospitalizations with equal weight. We apply our main specification with year, month, and day of the week of birth fixed effects, as well as controls (maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data). Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table C.8 Robustness: The Effect of Day of the Month of Birth on the SES and Health Indices, Sample of Survivors/Danish Residents in 2017, ITT.

	(1)	(2)	(3)	(4)
Panel A: SES Index				
Born 1-3	0.005	0.005	-0.002	0.009
	(0.012)	(0.012)	(0.011)	(0.014)
No. of obs.	76,468	76,468	75,911	17,401
Panel B: Health Index				
Born 1-3	0.031***	0.030***	0.029***	0.030**
	(0.011)	(0.011)	(0.011)	(0.014)
No. of obs.	77,299	77,299	76,722	17,593
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. The control variables added in columns (3) and (4) are maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.9 Placebo: The Effect of Day of the Month of Birth on the Probability of Emigration or Death in the Towns Sample (Outside Copenhagen), ITT.

	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, month, and day of the week of birth fixed effects, as well as controls (maternal age at birth and indicators for child sex, the child being born in a hospital, and the father not being observed in the administrative data). We do not control birth weight, as we do not observe nurse registrations in the sample. Instead, we add town fixed effects. Individuals enter the sample if observed in the administrative data at least once between 1977-2017. While deaths are measured in the death registry for the 1970-2020 period, individuals in our towns sample need to survive to 1977 to be observed in our register data at Statistics Denmark. Thus, for them we only observe deaths in the 1977-2020 period. This period deviates from the period observed for our nurse 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 covering the 1968 period and beyond). MDV: Mean of the dependent variable in the control group. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.10 Additional Heterogeneity: The Effect of Extended Nurse Care on the SES and Health Indices across Subgroups, ITT and IV Estimates.

	(3) Firstborn child		(4) Mother's age ≥ Median (24 yrs)		(5) Hospital birth		(6) Missing father		(7) Mother ≥ comp. edu.		(8) Breastfeeding at one month	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Panel A: SES Index												
Born 1-3	0.007	0.002	0.001	0.008	0.003	0.006	0.007	0.003	0.011	-0.005	0.013	-0.003
	(0.017)	(0.016)	(0.017)	(0.016)	(0.013)	(0.026)	(0.012)	(0.053)	(0.015)	(0.017)	(0.019)	(0.016)
P-value (Difference)	0.827		0.757		0.926		0.945		0.468		0.508	
Treatment Table	0.011	0.003	0.002	0.014	0.006	0.013	0.012	0.006	0.019	-0.010	0.022	-0.005
	(0.028)	(0.029)	(0.031)	(0.027)	(0.021)	(0.055)	(0.021)	(0.107)	(0.026)	(0.030)	(0.031)	(0.025)
MDV	-0.10	0.09	-0.07	0.07	0.06	-0.23	0.01	-0.29	-0.22	0.26	-0.11	0.11
No. of obs.	36,902	45,698	39,079	43,521	65,006	18,023	79,027	4,002	42,394	37,489	31,388	42,712
Panel B: Health Index												
Born 1-3	0.035**	0.021	0.016	0.039**	0.024*	0.038	0.031***	-0.030	0.027*	0.028*	0.022	0.035**
	(0.017)	(0.015)	(0.016)	(0.015)	(0.012)	(0.025)	(0.011)	(0.049)	(0.016)	(0.016)	(0.018)	(0.015)
P-value (Difference)	0.530		0.300		0.618		0.229		0.974		0.601	
Treatment Table	0.058**	0.040	0.029	0.064**	0.040*	0.079	0.054***	-0.060	0.046*	0.050*	0.037	0.055**
	(0.027)	(0.028)	(0.031)	(0.025)	(0.021)	(0.052)	(0.020)	(0.099)	(0.027)	(0.028)	(0.030)	(0.024)
MDV	-0.03	0.03	-0.02	0.02	0.03	-0.10	-0.00	0.01	-0.08	0.08	-0.03	0.04
No. of obs.	38,868	47,886	41,022	45,732	68,237	19,099	82,591	4,745	44,342	39,061	33,099	44,677

Notes: Each cell shows estimates from separate regressions estimated across subgroups of the population (denoted in the column headings). Regressions include year, month, and day of the week of birth fixed effects but exclude other control variables. 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.

Table C.11 Mechanisms: The Effect of Day of the Month of Birth on Child Being Observed in Three-Year Examination and a Childhood Good Health Index at Three and Six Years, ITT.

	Observed, Age 3	Health Index, Age 3	Health Index, Age 6
Panel A: Without Low BW Control			
Born 1-3	-0.034	0.230***	0.284***
	(0.025)	(0.079)	(0.073)
MDV	0.62	0.00	0.00
No. of obs.	4,369	1,971	1,861
Panel B: With Low BW Control			
Born 1-3	-0.034	0.220***	0.284***
	(0.025)	(0.079)	(0.073)
MDV	0.62	0.00	0.00
No. of obs.	4,369	1,971	1,861

Notes: Each column presents estimates from a separate regression of the respective outcome on an indicator for the child being born during the first three days of a month. The first column includes children who do not participate in the three year examination. The good health indices (at ages three and six, respectively) are equally weighted indices based on the underlying age-specific indicators for having received antibiotics, having been admitted to hospital, having been exposed to infectious disease, and having completed the recommended vaccinations. Both indices are based on standardized underlying variables following the procedure described in Section 3.2. The indices are set to missing if one or more underlying health measures are missing. MDV: Mean of the dependent variable in the control group. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.12 Mechanisms: The Effect of Day of the Month of Birth on Childhood Height, ITT.

	Height at				
	Age 3	Age 6	Age 7	Age 10	Age 13
Panel A: Without Low BW Control					
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
Panel B: With Low BW 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 during the first three days of a month for the sample of CPC children observed in the nurse records. All regressions for childhood height control for age at measurement. MDV: Mean of the dependent variable in the control group. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table C.13 Mechanisms: The Effect of Day of the Month of Birth on Indicators for Mother-Reported Full Uptake of Recommended Vaccinations and Indicators for Ever Having Used Antibiotics, ITT.

	Vacc., Age 3	Vacc., Age 6	Antib., Age 1-3	Antib., Age 3-6
Panel A: Without Low BW Control				
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
Panel B: With Low BW 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

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

Table C.14 Mechanisms: The Effect of Day of the Month of Birth on Child Ever Having Been Admitted to Hospital and of Ever Having Been Diagnosed with Infectious Disease (Mother-Reported), ITT.

	Hosp, 1-3	Hosp, 3-6	Inf. Disease, 1-3	Inf. Disease, 3-6
Panel A: Without Low BW Control				
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
Panel B: With Low BW 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

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

Table C.15 Mechanisms: The Effect of Day of the Month of Birth on Developmental Milestones Indices (Mother-Reported Age at Completion), ITT.

	Language	Motor	Eating	Dressing	Soc. Inter.	Toilet	Age 3, Avg.
Panel A: Without Low BW Control							
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
Panel B: With Low BW 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

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

D Family Spillovers and Intergenerational Impacts

Table D.1 The Effect of Day of the Month of Birth on Mothers' Fertility and Labor Market Outcomes, ITT.

	(1)	(2)	(3)	(4)
Panel A: Age at Birth				
Born 1-3	0.044	0.074	0.055	-0.031
	(0.079)	(0.078)	(0.079)	(0.099)
No. of obs.	47,886	47,886	47,622	10,925
Panel B: Spacing (Days)				
Born 1-3	2.813	-0.061	2.279	-0.142
	(16.519)	(16.464)	(16.682)	(20.687)
No. of obs.	36,565	36,565	36,342	8,287
Panel C: No. of Children				
Born 1-3	-0.012	-0.011	-0.008	0.006
	(0.014)	(0.014)	(0.014)	(0.018)
No. of obs.	47,886	47,886	47,622	10,925
Panel D: Average Earnings 40-65 (2015 DKK)				
Born 1-3	-184.422	-1065.816	-1059.309	-1431.780
	(1841.070)	(1819.180)	(1834.339)	(2319.308)
No. of obs.	47,062	47,062	46,807	10,726
Panel E: Average Employment 40-65				
Born 1-3	-0.003	-0.005	-0.005	-0.008
	(0.005)	(0.005)	(0.005)	(0.006)
No. of obs.	46,510	46,510	46,304	10,612
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression for mothers with their first child in the nurse records. Control variables are indicators for focal child sex, low birth weight status, child being born in Copenhagen, child being born in a hospital, child being first-born, and the father not being observed in the administrative data. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table D.2 Family Spillovers: The Effect of Day of the Month of Birth on the SES and Health Indices (ITT), Sample of First Children of each Family in the Nurse Records.

	(1)	(2)	(3)	(4)
Panel A: SES Index				
Born 1-3	0.019	0.015	0.010	0.018
	(0.013)	(0.013)	(0.013)	(0.017)
No. of obs.	61,595	61,595	61,438	14,089
Panel B: Health Index				
Born 1-3	0.037***	0.034***	0.033**	0.032**
	(0.013)	(0.013)	(0.013)	(0.016)
No. of obs.	64,603	64,603	64,246	14,719
YOB, MOB, and DOW FE		✓	✓	✓
Pre-treatment Controls			✓	✓
Only Born 1-7				✓

Notes: Each cell shows estimates from a separate regression. The sample only includes first trial-cohort children, i.e., excludes younger siblings. In column (2), we add year, month, and day of the week of birth fixed effects. The control variables added in columns (3) and (4) are maternal age at birth and indicators for child sex, child low birth weight status, child firstborn status, the child being born in Copenhagen, the child being born in a hospital, and the father not being observed in the administrative data. Robust standard errors are in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table D.3 Family Spillovers: The Effect of Day of the Month of Birth of a Focal Child on Sibling Outcomes, ITT.

	(1)	(2)	(3)	(4)	(5)	(6)
	Older Child Treated			Younger Child Treated		
Spacing	All	≤2yrs	>2yrs	All	≤2yrs	>2yrs
Panel A: Sibling SES Index						
Sibling Born 1-3	0.046**	0.061**	0.037	0.023	-0.012	0.042
	(0.018)	(0.030)	(0.023)	(0.020)	(0.032)	(0.025)
MDV	0.078	-0.025	0.137	-0.021	-0.034	-0.012
No. of obs.	33,738	12,383	21,355	26,934	10,825	16,109
Panel B: Sibling Health Index						
Sibling Born 1-3	0.012	0.040	-0.005	-0.009	-0.043	0.013
	(0.017)	(0.027)	(0.021)	(0.020)	(0.032)	(0.027)
MDV	0.086	0.023	0.122	-0.103	-0.040	-0.146
No. of obs.	35,637	13,134	22,503	28,349	11,436	16,913
Panel C: Sibling has a Nurse Record						
Sibling Born 1-3	0.011	0.013	0.009	0.002	0.013	-0.010
	(0.010)	(0.013)	(0.015)	(0.008)	(0.011)	(0.013)
MDV	0.703	0.747	0.657	0.871	0.883	0.858
No. of obs.	22,164	11,297	10,867	17,891	9,555	8,336
Panel D: Sibling Nurse Record is Empty for the 9 Months Visit						
Sibling Born 1-3	-0.026**	-0.037**	-0.014	-0.002	0.008	-0.013
	(0.012)	(0.016)	(0.018)	(0.011)	(0.016)	(0.016)
MDV	0.254	0.259	0.248	0.198	0.205	0.190
No. of obs.	14,501	7,817	6,684	14,409	7,771	6,638
Panel E: Sibling was Breastfed at One Month						
Sibling Born 1-3	0.024*	0.044**	-0.000	-0.001	-0.013	0.010
	(0.014)	(0.020)	(0.021)	(0.014)	(0.019)	(0.020)
MDV	0.538	0.531	0.547	0.649	0.639	0.660
No. of obs.	13,334	7,143	6,191	13,233	7,087	6,146

Notes: Each cell shows estimates from a separate regression of the outcome measured for the relevant sibling on an indicator for the focal child being born on the first three days of the month and fixed effects for sibling and focal child year of birth. Columns (1) and (4) use the full sample of sibling pairs, while the remaining columns focus on siblings with narrow or wider spacing. For further details on constraints in the data and the construction of the sibling samples, consult Section 3.2. MDV: Mean of the dependent variable in the control group for the relevant subgroup. Robust standard errors are in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table D.4 The Effect of Day of the Month of Birth on Fertility of Focal Children, Heterogeneity by Sex and Initial Disadvantage of the Focal Child, ITT.

Outcome	Childless		No. of Children		Age at First Child	
Panel A: First Gen. Sex	M	F	M	F	M	F
Born 1-3	0.006	-0.008	-0.040**	0.005	0.012	-0.006
	(0.007)	(0.006)	(0.017)	(0.016)	(0.106)	(0.094)
MDV	0.26	0.17	2.12	2.11	29.21	26.53
No. of obs.	44,247	42,286	32,793	35,200	32,787	35,187
Panel B: First Gen. Initial Disadv.	No	Yes	No	Yes	No	Yes
Born 1-3	0.007	-0.015**	-0.023	-0.006	-0.020	-0.018
	(0.006)	(0.007)	(0.015)	(0.019)	(0.094)	(0.113)
MDV	0.21	0.23	2.11	2.14	28.28	27.15
No. of obs.	51,233	35,300	40,605	27,388	40,596	27,378

Notes: Each cell presents the estimate of a separate regression. We divide the sample into female and male focal children and by the child's initial health status, respectively, in Panels (A) and (B). MDV: Mean of the dependent variable in the control group for the relevant subgroup. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table D.5 Placebo: The Effect of Day of the Month of Birth on Birth Outcomes in the Second Generation, Sample of Towns outside Copenhagen. Heterogeneity by Focal Child Initial Disadvantage, ITT.

	Focal Child Initial Disadvantage					
	No	Yes	No	Yes	No	Yes
Second Gen. Outcome	Birth Weight (g)		Low BW		Preterm	
Born 1-3	4.844	2.817	-0.001	0.002	0.000	0.002
	(8.070)	(10.177)	(0.003)	(0.004)	(0.003)	(0.004)
MDV	3469.73	3459.52	0.05	0.05	0.06	0.06
No. of obs.	59,407	39,407	59,407	39,407	58,796	38,993

Notes: Each cell shows estimates from separate regressions (for second generation birth outcomes) across subgroups of the population (defined by the initial disadvantage status of the focal individual, who is now a parent). The initial disadvantage status indicator is one if the focal child was born in a hospital, was born to a young mother, and/or has no father registration. The samples consists of all children born to either focal mothers or fathers (born 1959-1967), respectively, in the three major towns outside Copenhagen. If both parents are in the sample, we keep the mother spell. The estimates come from our heterogeneity specification including fixed effects for focal children's year, month, 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. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Table D.6 The Effect of Day of the Month of Birth on Birth Outcomes in the Second Generation, Heterogeneity by Focal Child Initial Disadvantage Status, Results for Separate Samples of Focal Mothers/Fathers, ITT.

	Focal Child Initial Disadvantage					
	No	Yes	No	Yes	No	Yes
Second Gen. Outcome	Birth Weight (g)		Low BW		Preterm	
Panel A: All Children of Focal Mothers						
Born 1-3	-14.912	8.084	0.003	-0.005	0.009**	-0.003
	(9.995)	(11.790)	(0.004)	(0.005)	(0.004)	(0.005)
MDV	3440.25	3362.27	0.05	0.07	0.06	0.07
No. of obs.	41,169	29,410	41,169	29,410	40,578	28,910
Panel B: All Children of Focal Fathers						
Born 1-3	-7.754	19.617	-0.001	-0.005	0.003	0.000
	(10.039)	(12.601)	(0.004)	(0.005)	(0.004)	(0.005)
MDV	3458.56	3403.52	0.05	0.06	0.06	0.07
No. of obs.	39,769	26,465	39,769	26,465	39,356	26,150

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

Table D.7 The Effect of Day of the Month of Birth on Educational Attainment in the Second Generation (Second Generation Children Born Prior to 1995). Heterogeneity by Focal Child Initial Disadvantage Status, ITT

Second Gen. Outcome	Above Compulsory Education	
First Gen. Initial Disadv.	No	Yes
Born 1-3	-0.017***	0.013*
	(0.006)	(0.008)
MDV	0.81	0.75
No. of obs.	45,874	33,885

Notes: Each cell shows estimates from separate regressions (for second generation educational outcomes) across subgroups of the population (defined by initial disadvantage status of the focal individual, who is now a parent). The samples consists of all children born to focal individuals prior to or in 1995. If a child has both parents (focal children) in our data, we keep the mother spell. The estimates come from our heterogeneity specification including fixed effects for focal children's year, month and day of the week of birth fixed effects, as well as focal child sex. MDV is the mean of the dependent variable in the control group for the relevant subgroup. Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

E Transcription Methods and Validity

E.1 Detection and Classification of the Treatment Table

To detect whether a nurse record contains the treatment table, we use the ML layout classification procedure described in Dahl et al. (2023a). We summarize the method and results here. Figure 1 shows a typical nurse record with a treatment table on the third page. As the page containing the treatment table varies, an approach that always considers the third page is not helpful. We use an unsupervised ML method to organize pages according to their appearance and layout such that similar looking pages are grouped. This approach is useful in out setting as scan quality and alignment is very uniform across the collection of documents, and thus 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 use a transfer learning approach and re-use a VGG-16 neural network (Simonyan and Zisserman, 2015) pre-trained to perform classification on ImageNet.[60] We strip the classification layers of the VGG-16 network and rely on the 512-dimensional vector representation that can be extracted from the final convolutional layer of the network after pooling. 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, we cluster all feature vectors 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 in total, and assigning a label to each of them. In this process, we use t-distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008) to visualize the feature vectors and clusters. The t-SNE method produces embeddings of the feature vectors in two-dimensional space that

[60]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`.

preserves their structure, i.e., points that are close in feature space also tend to be close in the low-dimensional embedding space. Appendix Figure E.1 shows the results with a clear separation of the different clusters, and thus the different page types.

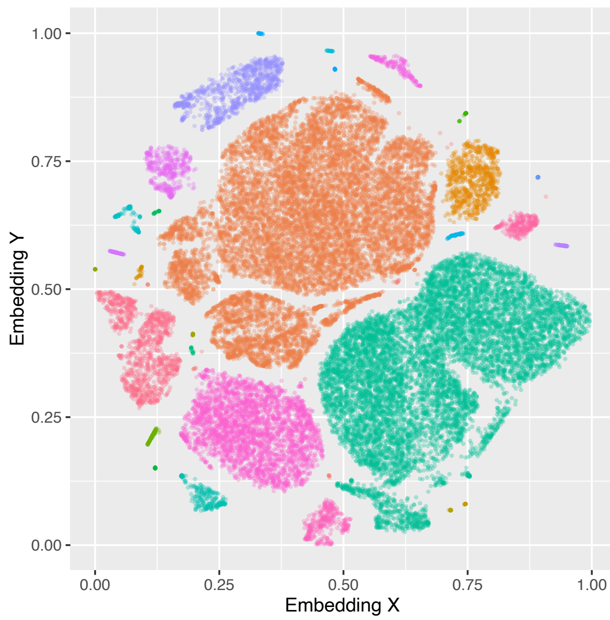


Fig. E.1 2D t-SNE Visualization of the Feature Space of the Record Pages.

Notes: Each point corresponds to the embedding of a record page and its color shows the cluster it was assigned to by the clustering algorithm. There are a total of 37 page clusters that we manually annotated; four of these contain the treatment pages. For the figure, we subsampled the points to reduce cluttering, so the figure only displays 30,000 randomly sampled page embeddings.

We evaluate the clustering procedure on a test dataset of 4,000 randomly selected records (manually reviewed to find the page with a treatment table). We compare the predictions from the clustering approach to these manually identified treatment tables. Appendix Table E.1 reports the validation results, showing no errors on the test dataset, meaning that the clustering procedure correctly classified all 4,000 records. This validates the ability of the clustering procedure to identify pages, and thus records, containing a treatment table.

We next classify the cells of the treatment table to determine if they are filled out or empty, and in turn, use these classifications to identify the actual provision of treatment as registered by the nurses. We use a supervised convolutional neural network that we trained on 2,000 manually transcribed random cells. We classify the cells into three classes: hand written cells, empty cells, and machine written cells, where the machine written cells would

Table E.1 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 4,000 records (10,914 pages). The treatment detection model does not rely on any segmentation, but instead detects the presence of the whole page containing the treatment table.

indicate that a cell should not be filled in by the nurses. We test the accuracy of this procedure on a held-out test set and find that it achieves an accuracy of 99.4%.

E.2 Detection and Transcription of Other Fields in the Nurse Records

We next transcribe fields (i.e., the cells of a table) containing nurse registrations for first-year visits in the family home. Page one of Figure 1 shows an example of a page with information on first-year visits. We use an approach similar to Dahl et al. (2023a). We start on the first page of each journal and try to cut the fields into separate images. This fails if fields are missing, and in such cases, we consider subsequent pages of the record until we get a hit.

We cut the fields into separate images using point set registration, which aligns a set of points between an image and a template (Besl and McKay, 1992). Dahl et al. (2023a) apply this method in a similar setting. We exploit the tabular structure of the nurse records to extract a set of points containing the intersections between the vertical and horizontal lines of the tables. To extract the lines, we use a pre-trained UNET model for semantic segmentation (Dahl and Westermann, 2023). After extracting the set of intersection points, we align the points with a pre-specified table template using the robust and efficient probabilistic point-set registration (FilterReg) of Gao and Tedrake (2019). This gives us a transformation matrix that we use together with the template to crop each field into a separate image.

To transcribe the segmented fields, we use neural networks trained on a manually transcribed subset of the nurse records.[61] The exact network design varies by the field we tran-

[61]Part of the manual transcriptions were done by Andersen et al. (2012) and Bjerregaard et al. (2014),

scribe, but it always derives from two main architectures: A convolutional neural network (CNN) similar to the date transcription models of Dahl et al. (2022), or a vision transformer (ViT) neural network (Dosovitskiy et al., 2021).

The convolutional neural network is a slightly modified EfficientNetV2-S model (Tan and Le, 2021). Tan and Le (2021) improve the performance of the EfficientNet model introduced in Tan and Le (2019) by using training-aware neural architecture search and scaling to jointly optimize training speed and parameter efficiency. We replace its final layer by a series of classification heads, each corresponding to an element in the sequence of the field being transcribed, inspired by Goodfellow et al. (2013); this is similar to the strategy in Dahl et al. (2022), but with a broader set of classification heads, as we are not exclusively transcribing dates. We employ an optimization procedure similar to Dahl et al. (2022) and likewise transfer learn from a model pretrained on ImageNet21k (Russakovsky et al., 2015).[62] This means we use SGD with momentum as the optimizer and use dropout (Srivastava et al., 2014), stochastic depth (Huang et al., 2016), label smoothing (Szegedy et al., 2016), adaptive gradient clipping (Brock et al., 2021), RandAugment (Cubuk et al., 2020), and random erase (Zhong et al., 2020). See Appendix Table E.2 for additional details on the hyperparameters.

The vision transformer we use is a modified DeiT III (base) model (Touvron et al., 2022). Touvron et al. (2022) improve the performance of the DeiT network introduced in Touvron et al. (2021a) by building upon work by Wightman et al. (2021), who improve the performance of ResNets (He et al., 2016) by using modern training techniques. To improve the training of ViTs, Touvron et al. (2022) employ techniques such as stochastic depth (Huang et al., 2016) and LayerScale (Touvron et al., 2021b). The overall strategy we employ is to replace the model's final layer by a decoder transformer. We use a three-layer decoder transformer with eight heads and feedforward network dimension of 512. The decoder treats the output of its DeiT III-feature extractor as a sequence and then decodes it to produce a transcription of the image's content. We mostly follow the training procedure of Touvron et al. (2022),

leaving only minor need for additional labelling.

[62]With the exception of CNNs for transcription of nurse names; for these, we first train a model on the database of names from Dahl et al. (2023b) and then use it to transfer learn from.

specifically their procedure for finetuning their ViT-B model trained on ImageNet21k to ImageNet1k (Russakovsky et al., 2015), and likewise transfer learn from this model.[63] This means we use the LAMB optimizer (You et al., 2019) and label smoothing (Szegedy et al., 2016). We deviate in one aspect, as we use RandAugment (Cubuk et al., 2020) for image augmentation rather than the 3-Augment method introduced in Touvron et al. (2022). See Appendix Table E.2 for additional details on the hyperparameters.

We train our networks using `timmsn`, a Python package for transcription of text from images, including handwritten text recognition. This module uses the PyTorch Image Models library (Wightman, 2019). Code with a full list of the exact architectures and training parameters of all models, including code to replicate our transcription results, is available upon request and will be made available at `https://github.com/TorbenSDJohansen/cihvr-transcription`.

While the architecture and training procedure always follow Appendix Table E.2 (CNN or ViT), we train a series of models with different final layer, as the fields of the nurse records differ with respect to length, image size (including aspect ratio), and 'alphabet" (that is, the set of characters that can occur in the field). To account for these differences, we group fields that are similar with respect to length and alphabet.[64] Appendix Table E.3 shows the resulting ten groups with the number of assigned fields, maximum length of the group contents, and alphabet. Most groups, and thus fields, consist of sequences of numbers. This motivates our use of not only models trained on any one group, but also models trained on unions of some of the groups.

We train neural networks for each of the ten groups of Appendix Table E.3. We also train a CNN and a ViT for the union of breastfeeding at seven days old (breastfeed-7-do) and born prior to the due date of the pregnancy (birth prior to due date), as both groups consist of fields where the content is a circle around a number; we label these models "Circle". Finally, we

[63]With the exception of ViTs for transcription of nurse names; for these, we first train a model on the database of names of Dahl et al. (2023b) and then afterwards use it to transfer learn from.

[64]For convenience, we treat nurse first and last names as separate fields, even though both the first and last name of a given nurse is contained within the same field in the records.

Table E.2 Transcription Model Parameters.

	CNN	**ViT**
Feature extractor	EfficientNetV2-S	ViT-B (DeiT III)
Decoder	Linear	Transformer
RandAugment	✓	✓
Gradient clip	✓	✓
Dropout prob.	0.4	-
Stochastic depth prob.	0.25	0.15
Epochs	250	250
Learning rate	0.5	0.001
Optimizer	SGD	LAMB
Random erase prob.	0.4	-
Label smoothing	0.1	0.1
Weight decay	7e-06	0.02
Warmup epochs	10	5
LR decay	Cosine	Cosine

Notes: The table shows hyperparameters of our two transcription neural network architectures, including parameters related to model training and not directly to architecture. The parameters for the CNN architecture are taken from Dahl et al. (2022) (with only very minor differences). The parameters for the ViT architecture are inspired by Touvron et al. (2022), specifically from their finetuning on ImageNet1k of their ViT-B model pre-trained on ImageNet-21k, with the main differences being training for more epochs and using RandAugment for image augmentation rather than their proposed 3-Augment (to match number of epochs and augmentation method of our models with CNN architecture), as well as replacing the final layer with a 3-layer transformer decoder model. The specific models vary slightly due to differences in fields used for training; specifically, the image resolution, sequence length, alphabet, and batch size (as a result of differences in resolution and availability of 1 vs 2 GPUs for training) vary. See Appendix Table E.4 for a list of differences across our models. Code with a full list of the exact architectures and training parameters of all models, including code to replicate our transcription results, is available upon request and will be made available at a later date at `https://github.com/TorbenSDJohansen/cihvr-transcription`

train a ViT for the union of duration of breastfeeding (dura-any-breastfeed), length (length), number of weeks born prior to due date (weeks prior to due date), Table B visit information (tab-b), and weight (weight), as all groups consist of sequences of one to five integers; we label this model "Integer seq.". This leads to a total of 23 neural networks (11 CNNs and 12 ViTs). Appendix Table E.4 shows the image resolution, sequence length, and batch size of the ViT (Panel A) and CNN (Panel B) models for each field.[65]

Having trained these 23 neural networks, we next select which ones to use for transcription. We evaluate each model on a hold-out (that is, not used for training of the model) test set consisting of images from the same field group that the model was trained on. This means that we evaluate a model on the same field group it was trained on. For models trained

[65] A further difference has to do with the alphabet of the models. These are defined as the union of the alphabets of the different groups of fields used for the specific model (see Appendix Table E.3), plus some other special characters to denote, e.g., beginning of sequence, end of sequence, and separator.

Table E.3 Grouping of Fields from the Nurse Records.

Group	#Fields	Maximum Length	Alphabet
breastfeed-7-do	1	1	$\{1, 2, 3\}^*$
dura-any-breastfeed	1	2	$\{0, 1, \ldots, 9\}$
date	7	4	$\{0, 1, \ldots, 9\}$
length	2	3	$\{0, 1, \ldots, 9\}$
birth prior to due date	1	1	$\{1, 2\}^*$
no. of weeks prior to due date	1	2	$\{0, 1, \ldots, 9\}$
tab-b	112	2	$\{0, 1, \ldots, 9\}$
weight	8	5	$\{0, 1, \ldots, 9\}$
nurse-name (first)	3	k^{**}	$\{a, b, \ldots, \text{å}\}$
nurse-name (last)	3	k^{**}	$\{a, b, \ldots, \text{å}\}$

Notes: The table shows fields of the nurse records grouped together in such a way that fields which are similar with respect to sequence length and alphabet of their contents are put into one group. The first column refers to the name given to the specific group of fields. The second column shows the number of fields of the given group. The third column shows the maximum length of the content of any field of the given group. The fourth column shows the alphabet of the fields of the given group. *The alphabet of these fields are specifically a *circle* being put around one of the digits shown, the digits being pre-printed on the records. **While there is no clear limit to the length of a name, the longest name in our training dataset contains 14 characters.

on multiple groups of fields (e.g., the “Integer seq.”-model), we evaluate them on all groups separately to avoid selecting a model that is worse for a specific group, but reaches a better transcription accuracy on the union of groups it was trained on. Appendix Table E.5 shows the models we selected by choosing those with the highest full sequence accuracy on their corresponding test sets.

Table E.4 Model Differences.

Model	**Fields/Groups (see Table E.3)**	**Resolution**	**Seq. len.**	**Batch size**
	Panel A: ViT-based models			
BF 7 do.	breastfeed-7-do	117x537	4	256
Circle	breastfeed-7-do, preterm-birth	100x350	4	256
Dura. any BF	dura-any-breastfeed	88x284	4	512
Date	date	67x181	7	1024
Length	length	109x297	4	512
Birth prior to due date (Y/N)	birth prior to due date	107x249	4	512
Weeks prior due date	weeks prior to due date	100x193	4	512
Integer	tab-b	79x121	4	512
Weight	weight	80x258	7	512
First name	nurse-name (first)	91x530	20	512
Last name	nurse-name (last)	91x530	20	512
Integer seq.	dura-any-breastfeed, length, preterm-birth-weeks, tab-b, weight	90x230	7	1024
	Panel B: CNN-based models			
BF 7 do.	breastfeed-7-do	117x537	2	256
Circle	breastfeed-7-do, preterm-birth	100x350	2	256
Dura. any BF	dura-any-breastfeed	88x284	2	512
Date	date	67x181	3	1024
Length	length	109x297	2	512
Birth prior to due date (Y/N)	birth prior to due date	107x249	2	512
Weeks prior to due date	weeks prior to due date	100x193	2	512
Integer	tab-b	79x121	2	1024
Weight	weight	80x258	5	1024
First name	nurse-name (first)	91x530	18	256
Last name	nurse-name (last)	91x530	18	256

Notes: The table shows differences of hyperparameters of the 23 neural networks, beyond those between the CNN- and the ViT-based models (see Appendix Table E.2 for differences between the CNN- and the ViT-based models). The differences are all related to the groups of fields used for training, as they differ in resolution (including aspect ratio) and sequence length, which, together with varying availability of 1 vs 2 GPUs, led to different batch sizes. Note how the sequence lengths of this table differ from those of Appendix Table E.3, often being longer. This is due to the addition of certain special characters such as beginning of sequence and end of sequence tokens being pre- and appended to the sequences, respectively, for some models.

Table E.5 Model Differences – Selected Models.

Model	Fields/Groups (see Table E.3)	Resolution	Seq. len.	Batch size
	Panel A: ViT-based models			
Circle	breastfeed-7-do, birth prior to due date	100x350	4	256
Integer seq.	dura-any-breastfeed, length, weeks prior to due date, tab-b, weight	90x230	7	1024
Last name	nurse-name	91x530	20	512
	Panel B: CNN-based models			
Date	date	67x181	3	1024
First name	nurse-name	91x530	18	256
Weight	weight	80x258	5	1024

Notes: The table shows differences of hyperparameters of the final six neural networks selected among those of Appendix Table E.4, beyond those between the CNN- and the ViT-based models (see Appendix Table E.2 for differences between the CNN- and the ViT-based models). The differences are all related to the groups of fields used for training, as they differ in resolution (including aspect ratio) and sequence length, which, together with varying availability of 1 vs 2 GPUs, led to different batch sizes. Note how the sequence lengths of this table differ from those of Appendix Table E.3, often being longer. This is due to the addition of certain special characters such as beginning of sequence and end of sequence tokens being pre- and appended to the sequences, respectively, for some models.

Performance of Transcription Appendix Table E.6 shows transcription performance across field groups of the nurse records. The groups are at a slightly more granular level than those of Appendix Table E.3, as the Table B group is now split into 16 groups, according to the 16 different columns of the group; each column represents seven different fields. We measure accuracy by the proportion of full sequences predicted with no errors. This means that a transcription of a name where 13 of the 14 letters are transcribed correctly will still be counted as an error. This accuracy measure is more conservative compared to the often reported character accuracy.[66] Appendix Table E.6 also reports performance metrics when we drop empty fields, showing that the high transcription accuracy is not driven by correctly predicting a large share of empty cells, which are "easy" to get right. The last column of the table reports the share of non-empty cells.[67]

[66] We do not evaluate on character accuracy, as we need the full sequence to be transcribed correctly. For example, while a weight of 6,000 grams is, in practical terms, significantly different from a weight of 5,000 grams, under the character accuracy metric, the latter would still be scored as 75 percent the same as the former (as all the zeros are transcribed correctly). To avoid this, we chose the sequence accuracy metric for our task.

[67] Note that this number *does not* represent the share of non-empty cells in the total collection of records for a given group, but rather just for the *test* set.

Table E.6 Automated Transcription Performance.

	Transcription accuracy (%)		Share non-empty (%)
	All	Non-empty	
Babbles	92.9	97.5	61.5
Breastfeeding 7 days	99.4	99.7	91.4
Care and cleanliness	96.7	99.0	60.8
Date	97.2	96.8	77.5
Duration breastfeeding	97.6	97.9	97.3
Home economic status	96.0	97.6	46.5
Home harmony	97.7	98.9	26.5
In air	91.7	97.5	61.2
Length	99.0	99.0	97.2
Lifts head	92.6	94.9	63.4
Mother daily hours working at home	88.8	99.5	60.7
Mother daily hours working outside home	93.0	99.5	70.8
Mother mental capacity	96.8	97.4	40.9
Mother physical capacity	97.1	98.1	41.6
Number of daily meals	74.5	74.0	72.9
Nursery or kindergarten	93.9	93.3	69.9
Nutrition	96.2	97.0	80.2
Own bed	96.0	99.7	59.7
Birth prior to due date	99.0	99.5	88.4
Weeks prior to due date	97.3	80.1*	12.5
Sits	91.9	91.8	70.3
Smiles	93.0	98.0	61.5
Weight	97.8	97.7	97.3
Nurse first name	95.2	93.6	57.0
Nurse last name	95.0	93.2	57.0

Notes: The table shows the accuracy (%) of the ML transcriptions for separate groups 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 on the full test sample. The third column shows the accuracy when excluding empty fields. The fourth column shows the share of observations of the test set that is non-empty for each group. *The low sequence accuracy for non-empty number of weeks prior to due date is due to inconsistencies regarding manual labelling of ranges such as "1-2". In those cases, the label might either say 1 or 2, meaning that it is not possible to do better than guessing one of the two numbers for a number of these cases. Allowing the number of weeks born prior to the due date to differ by one increases the sequence accuracy to 95.3% for the non-empty cases and to 99.2% for the full sample.

F The Coverage of the Copenhagen Nurse Records

In this section, we discuss the coverage of our nurse records in the 1959-1967 cohorts and assess whether the nurse program in Copenhagen was universal. As we do not observe the number of resident children in Copenhagen at the time in either aggregate or individual level data, we use aggregate data from yearbooks and data on place of birth registrations from individual-level register data at Statistics Denmark.

The Coverage of the Nurse Records Appendix Table F.1 shows 1959-1967 statistics compiled by Copenhagen officials (Copenhagen Statistical Office, various years) and figures from our nurse records.

Table F.1 Coverage of the Copenhagen Yearly Statistics on the Nurse Program and of the Nurse Records.

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

Notes: Columns two and three are based on aggregate statistics from the Copenhagen Statistical Yearbooks (Copenhagen Statistical Office, various years), column four comes from the nurse records (number of nurse records including unmatched records).

There are two main findings from the table: First, the number of transcribed nurse records corresponds nicely to the aggregate records on infants entering treatment in the years under consideration. In most years, however, we observe more infants in our nurse records than recorded in the yearbooks. The yearbooks do not contain descriptions of the data, making it hard to assess the reasons for this discrepancy. Those may include (i) that the yearly municipal count of records is incomplete (due to nurses not handing in the children's records in time for the count) or (ii) that officials did not include infants who moved to Copenhagen during the year. Second, the aggregate figures on discontinuation of supervision in the first-year program indicate that up to 30 percent of infants left the first-year program. This figure lends further credibility to our first stage results (treatment assigned at around one year).

Coverage of the Nurse Program Did the Copenhagen nurse program reach out universally to all residents? The aggregate figures in the statistical yearbooks contain data on individuals under supervision but no figures on the total number of eligible resident infants in Copenhagen at the time. The yearly reports state that 90 percent of Copenhagen-born children entered supervision (Copenhagen Statistical Office, various years).

We assess the coverage of the nurse program using our data: In our individual level administrative data, we observe the number of Copenhagen births, rather than Copenhagen resident status. 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.

Appendix Table F.2 shows data on all Copenhagen births in the 1959-1967 period from the administrative data. It displays the type of birth registrations for all Copenhagen births (100,854) and by match status with the nurse records. Of the individuals born in Copenhagen, we observe 71,596 (73 percent) in the records.[68] However, among the infants born in Copenhagen but not covered in the nurse data, 46 percent were born in a hospital in the city. Thus, it is likely that many of those children did not actually reside in Copenhagen, but that their mothers came to the capital to give birth.

Similarly, considering individuals in the nurse records irrespective of place of birth, 82 percent of them were born in Copenhagen. Appendix Figure F.1 verifies that the share of Copenhagen births in our sample is stable across days of the month. The remainder of infants

[68] Factoring in the records (those without a personal identifier), we 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 F.2 Type of Birth Registration for Copenhagen Births (Administrative Data), by Match Status with Nurse Records, percent.

	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
No. of obs.	100,854	29,258	71,596

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

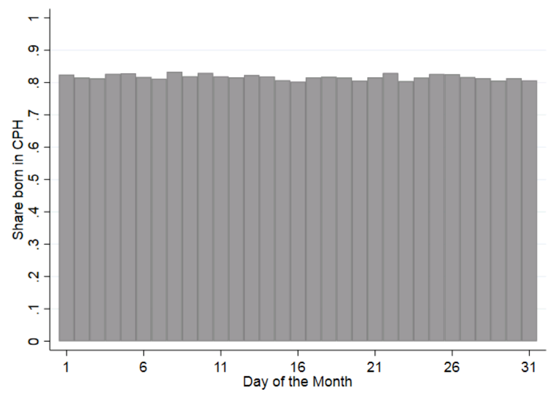


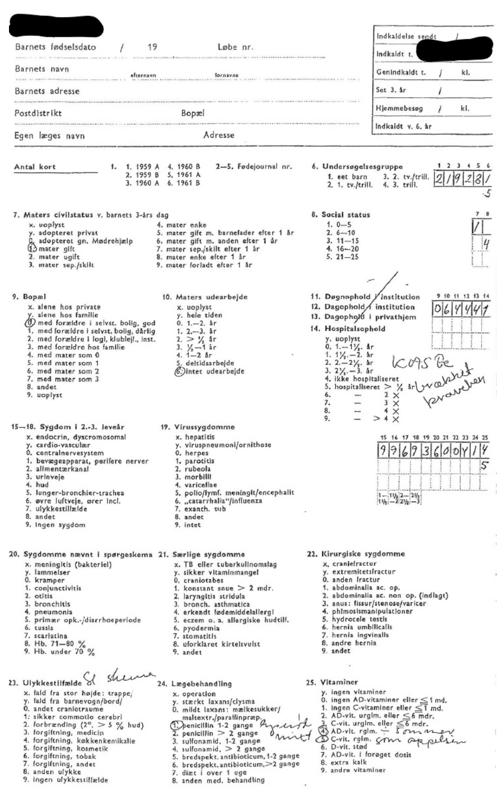
Fig. F.1 Share of Copenhagen Births in the Matched Nurse Records Across Days of the Month, 1959-1967.

in the nurse records were mainly born in adjacent municipalities (Frederiksberg, Tårnby, and Gentofte account for 57, 11, and 9 percent of births, respectively), suggesting that moves of new families played an important role. Thus, the individual level data on birth registrations support that the Copenhagen nurse program had universal outreach among residents.

G Data: The Copenhagen Perinatal Cohort (CPC)

The Copenhagen Perinatal Cohort (CPC) is a prospective cohort study of 9,125 infants born in the period 1959-1961 in Rigshospitalet, the largest hospital in Copenhagen. Focusing on subsets of (Copenhagen resident) children, we analyze outcome data that have been manually transcribed from CPC records (Merrick et al., 1983; Schack-Nielsen et al., 2010). Thus our choice of outcome variables is guided by (i) our assessment of relevant domains of childhood health, development, and parental health investments, and (ii) data availability.

The initial data collections in the CPC happened around birth during a perinatal interview, via registrations in the hospital records, and during medical examinations in the hospital on day one and five after birth. Follow-up data collections in the CPC were conducted at child age one, three, and six years. Mothers were invited to fill out a survey about their child's health and development and bring the child to a medical examination at Rigshospitalet at the given ages. Doctors transferred the mother reports into the CPC records. Additionally, the height of CPC children residing in Copenhagen was measured by school doctors around ages seven, ten, and 13 years. Due to mobility, there is considerable attrition over time of follow-up. For our analyses, we use data from follow-ups at ages three and six years (examinations and mother-reports transferred to the CPC records at the visits). Appendix Figure G.I shows a typical page from a CPC record for a child. The entries consist of a combination of filled check boxes and free text.



[redacted]

Barnets fødselsdato / 19 Løbe nr.

Barnets navn efternavn fornavne

Barnets adresse

Postdistrikt Bopæl

Egen læges navn Adresse

Indkaldelse sendt / [redacted]
Indkaldt t. [redacted]
Genindkaldt t. / kl.
Set 3. år /
Hjemmebesøg / kl.
Indkaldt v. 6. år

Antal kort

1. 1. 1959 A 4. 1960 B
2. 1959 B 5. 1961 A
3. 1960 A 6. 1961 B

2—5. Fødejournal nr.

6. Undersøgelsesgruppe
1. eet barn 3. 2. tv./trill.
2. 1. tv./trill. 4. 3. trill.

1 2 3 4 5 6
2 1 9 2 8 1
5

7. Maters civilstatus v. barnets 3-års dag
x. uoplyst
y. adopteret privat
0. adopteret gn. Mødrehjælp
1. mater gift
2. mater ugift
3. mater sep./skilt
4. mater enke
5. mater gift m. barnefader efter 1 år
6. mater gift m. anden efter 1 år
7. mater sep./skilt efter 1 år
8. mater enke efter 1 år
9. mater forladt efter 1 år

8. Social status
1. 0—5
2. 6—10
3. 11—15
4. 16—20
5. 21—25

7 8
1
4

9. Bopæl
x. alene hos private
y. alene hos familie
0. med forældre i selvst. bolig, god
1. med forældre i selvst. bolig, dårlig
2. med forældre i logi, klublejl., inst.
3. med forældre hos familie
4. med mater som 0
5. med mater som 1
6. med mater som 2
7. med mater som 3
8. andet
9. uoplyst

10. Maters udearbejde
x. uoplyst
y. hele tiden
0. 1.—2. år
1. 2.—3. år
2. > ½ år
3. ½—1 år
4. 1—2 år
5. deltidsarbejde
6. intet udearbejde

11. Døgnophold i institution
12. Dagophold i institution
13. Dagophold i privathjem
14. Hospitalsophold
y. uoplyst
0. 1.—1½. år
1. 1½.—2. år
2. 2.—2½. år
3. 2½.—3. år
4. ikke hospitaliseret
5. hospitaliseret > ½ år
6. — 2 ×
7. — 3 ×
8. — 4 ×
9. — > 4 ×

9 10 11 12 13 14
0 6 4 4 4 1

15—18. Sygdom i 2.-3. leveår
x. endocrin, dyscromosomal
y. cardio-vasculær
0. centralnervesystem
1. bevægeapparat, perifere nerver
2. alimentærkanal
3. urinveje
4. hud
5. lunger-bronchier-trachea
6. øvre luftveje, ører incl.
7. ulykkestilfælde
8. andet
9. ingen sygdom

19. Virussygdomme
x. hepatitis
y. viruspneumoni/ornithose
0. herpes
1. parotitis
2. rubeola
3. morbilli
4. varicellae
5. polio/lymf. meningit/encephalit
6. „catarrhalia"/influenza
7. exanth. sub
8. andet
9. intet

15 16 17 18 19 20 21 22 23 24 25
9 7 6 9 3 6 0 0 4 1 4
5

1— 1½ 2— 2½
1½ —2 2½ —3

20. Sygdomme nævnt i spørgeskema
x. meningitis (bakteriel)
y. lammelser
0. kramper
1. conjunctivitis
2. otitis
3. bronchitis
4. pneumonia
5. primær opk.-/diarrhoeperiode
6. tussis
7. scarlatina
8. Hb. 71—80 %
9. Hb. under 70 %

21. Særlige sygdomme
x. TB eller tuberkulinomslag
y. sikker vitaminmangel
0. craniotabes
1. konstant snue > 2 mdr.
2. laryngitis stridula
3. bronch. asthmatica
4. erkendt fødemiddelallergi
5. eczem o. a. allergiske hudtilf.
6. pyodermia
7. stomatitis
8. uforklaret kirtelsvulst
9. andet

22. Kirurgiske sygdomme
x. craniefractur
y. extremitetsfractur
0. anden fractur
1. abdominalia ac. op.
2. abdominalia ac. non op. (indlagt)
3. anus: fissur/stenose/varicer
4. phimosismanipulationer
5. hydrocele testis
6. hernia umbilicalis
7. hernia ingvinalis
8. andre hernia
9. andet

23. Ulykkestilfælde
x. fald fra stor højde: trappe/
y. fald fra barnevogn/bord/
0. andet cranietraume
1. sikker commotio cerebri
2. forbrænding (2°, > 5 % hud)
3. forgiftning, medicin
4. forgiftning, køkkenkemikalie
5. forgiftning, kosmetik
6. forgiftning, tobak
7. forgiftning, andet
8. anden ulykke
9. ingen ulykkestilfælde

24. Lægebehandling
x. operation
y. stærkt laxans/clysma
0. mildt laxans: mælkesukker/ maltextr./paraffinpræp.
1. penicillin 1-2 gange
2. penicillin > 2 gange
3. sulfonamid, 1-2 gange
4. sulfonamid, > 2 gange
5. bredspekt. antibioticum, 1-2 gange
6. bredspekt. antibioticum, >2 gange
7. diæt i over 1 uge
8. anden med. behandling

25. Vitaminer
y. ingen vitaminer
0. ingen AD-vitaminer eller ≤ 1 md.
1. ingen C-vitaminer eller ≤ 1 md.
2. AD-vit. urglm. eller ≤ 6 mdr.
3. C-vit. urglm. eller ≤ 6 mdr.
4. AD-vit. rglm.
5. C-vit. rglm.
6. D-vit. stød
7. AD-vit. i forøget dosis
8. extra kalk
9. andre vitaminer

Fig. G.I Excerpt from a CPC Examination Record for a Child at Age Three.

Appendix Table G.I summarizes the topics covered in the CPC records at ages three and six years. Due missing data, we cannot make use of all information. In particular, we do not have sufficient information on childcare attendance.

Table G.I CPC Data: Content of CPC Records.

Topic	Example Items	Age of Child 3 years	6 years
Background	Family background, childcare type	✓	✓
Health care take-up	Ever hospitalized and number of nights, treatment with specific drugs, vaccinations	✓	✓
Diagnoses	Infections, specific diagnoses (skin, lungs, ears)	✓	✓
Child health indicators	Height, weight	✓	✓
Child development	Developmental milestones	✓	

Notes: The table shows topics and items covered in the CPC records from mother survey reports (milestones) and physician examinations.

Thus, our main measures from the CPC analyses fall in three groups. First, we study mother-reported measures of health and health investments: indicators for the child having been hospitalized, having been exposed to infectious diseases (otitis, bronchitis, pneumonia), vaccination compliance, and antibiotics consumption (reported at ages three and six). We measure vaccine uptake as having all vaccinations in the standard vaccination program at the time (tetanus, polio, pertussis, tuberculosis, and small pox). We measure antibiotics consumption as ever having been treated with antibiotics in the relevant age span (penicillin or sulfonamide). Second, we consider physician-measured height at ages three and six years. Height is also recorded by school doctors (for Copenhagen residents) at ages seven, ten, and 13 years. We age-adjust our height analyses by linearly controlling for exact age at measurement in months. At the CPC examinations at ages three and six, physicians could opt to either tick a box for a height range or register exact height, resulting in many missing values among examined children for exact height. Third, parallel to our analyses on long-run outcomes, we construct indices of mother-reported good child health and developmental milestones. For the health indices at ages three and six years, we use information on hospital admissions, infections, vaccines, and antibiotics consumption. We compute a good health index equivalent to our long-run health index to increase statistical power. This index is missing if one or more of the underlying measures are missing. For the developmental milestones index, we aggregate information from 20 questions on child abilities across six domains (language, motor development, eating, dressing, social interactions, and toilet training). In each domain, mothers reported the age at which the child was able to perform the tasks. We create standardized scores for each domain, as well as an average developmental milestone index across all domains Flensborg-Madsen and Mortensen (2018): We transform the age for each of the 20 questions into standardized scores (Score $= \frac{\text{age} - \text{mean age of control group}}{\text{standard deviation of control group}}$). If a child had not yet achieved a milestone, we set the age to 36 months. Next, we average across these scores in six domains and for an overall developmental score. Finally, we standardize to arrive at the index used in our analyses.

H Data: Administrative Health Data (ICD Codes)

To create our good adult health index, we use data from the Danish National Patient Register on hospital admissions and diagnoses in the 1977-2018 period (Appendix Table H.I). In 1995, the diagnoses scheme changed from the ICD 8 to 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 (as diabetes is an absorbing state and as prevalence increase with age, we think this decision is reasonable). For mental health conditions, we measure whether an individual is ever observed in the discharge data with any mental health diagnosis.

Table H.I Hospital Diagnoses Defined in the Data and Underlying ICD Codes

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; DC00-DC97
Infections	000-136; DA00-DA99, DB00-DB99