RUHR
ECONOMIC PAPERS

Claudia Andreella
Martin Karlsson
Therese Nilsson
Matthias Westphal

The Long Shadows of Past Insults

Intergenerational Transmission of Health over 130 Years

#571

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, e-mail: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Roland Döhrn, Prof. Dr. Manuel Frondel, Prof. Dr. Jochen Kluve
RWI, Phone: +49 (0) 201/81 49-213, e-mail: presse@rwi-essen.de

Editorial Office

Sabine Weiler
RWI, Phone: +49 (0) 201/81 49-213, e-mail: sabine.weiler@rwi-essen.de

Ruhr Economic Papers #571

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-657-4

Ruhr Economic Papers #571

Claudia Andreella, Martin Karlsson, Therese Nilsson,
and Matthias Westphal

The Long Shadows of Past Insults

Intergenerational Transmission of Health over 130 Years

Ruhr Graduate School in Economics

UNIVERSITÄT DUISBURG ESSEN

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788657
ISSN 1864-4872 (online)
ISBN 978-3-86788-657-4

Claudia Andreella, Martin Karlsson, Therese Nilsson,
and Matthias Westphal[1]

The Long Shadows of Past Insults

Intergenerational Transmission of Health over 130 Years

Abstract

This paper investigates the intergenerational transmission of health in the very long run. Using a unique purpose-built administrative dataset on individuals born in Sweden between 1930–34 and their parents, we study the intergenerational transmission (IGT) of health and the impact of previous generations' health shocks on socioeconomic outcomes. Our results provide strong evidence in favour of IGT of health, in particular for males. In contrast to the existing literature that focuses on early life health outcomes, our paper shows that the effect on later-life mortality might be even more relevant. However, the story appears to be complex and multi-faceted: the IGT exhibits an inverted socioeconomic gradient, and the impact on socioeconomic outcomes is often very different from the effect on health.

JEL Classification: I12, J13

Keywords: Early environment; intergenerational transmission; Barker hypothesis; maternal health; infant health; socioeconomic status

July 2015

1 Claudia Andreella and Martin Karlsson, both UDE and CINCH; Therese Nilsson, CINCH; Matthias Westphal, UDE, CINCH, and RGS Econ. *– We are grateful to Ainhoa Aparicio, Tommy Bengtsson, Sonia Bhalotra, Gabriella Conti, Paul Devereux, Alain Gagnon, Matthew Lindquist and Mikko Myrskylä for valuable comments. We would also like to thank the participants of the SOLE/EALE Joint Conference in Montreal, the IUSSP seminar on Earlylife Determinants of Late-life Employment, Ill Health and Early Death in Lund, the Essen Health Conference 'Rethinking Barker', and seminar participants at the University of Duisburg-Essen. Financial support provided by the Center for Economic Demography at Lund University and the Crafoord Foundation is gratefully acknowledged.–* *All correspondence to: Claudia Andreella, CINCH-Health Economics Research Center, Edmund-Körner-Platz 2, 45127 Essen, Germany, e-mail: claudia.andreella@uni-due.de*

1 Introduction

The origins of societal inequality in health and socioeconomic outcomes are not well understood – but their persistence over time is quite remarkable. A large body of literature in economics measures intergenerational correlations in socioeconomic status. In the U.S., the elasticity of sons' earnings with respect to their fathers' earnings is 0.5-0.6 (Mazumder, 2005); whereas the corresponding number for Sweden is 0.25-0.3 (Jantti et al., 2006, Lindahl et al., 2012). Similar results are found for other socioeconomic outcomes such as IQ and educational attainment (Hertz et al., 2007). A nascent strand of the literature suggests that the persistence in social hierarchies may be even stronger than what a narrow focus on labour market outcomes would suggest (Clark, 2012).

A parallel, but considerably smaller, literature in health economics studies the intergenerational transmission of health. A typical indicator is birth weight, which has been shown to be persistent across generations. For American twin pairs, Royer (2009) finds that a 100-gram increase in birth weight associates with an 18-gram increase in the following generation; estimates including mother fixed effects are somewhat smaller but strongly significant. Intergenerational persistence is also observed in other health outcomes, such as longevity and self-assessed health (Trannoy et al., 2010) or body mass index (Classen, 2010).

Despite these efforts, the literature does not deliver much evidence of the extent to which the associations are modifiable. Comparing biological and adopted children, Thompson (2014) estimates the genetic component in the intergenerational correlation of some common chronic diseases and concludes that it is surprisingly low, explaining only 20-30% of the observed intergenerational link. This result, however, still begs the question as to whether the remaining intergenerational link is modifiable. A large and growing literature documents that early life shocks may have strong effects on health and socioeconomic outcomes in adulthood (Almond, 2006, Almond et al., 2009, Scholte et al., 2012). The same appears to hold for policy interventions (Aizer and Currie, 2014, Bhalotra et al., 2015, Bhalotra and Venkataramani, 2011, Bharadwaj et al., 2011). However, it is still largely unclear whether these early life influences contribute to reducing inequalities in health and socioeconomic outcomes in later generations. Increasing our knowledge about this issue would clearly be highly desirable, since it may be the case that public interventions generate returns over a very long time period.

It is the purpose of this paper to closely examine the intergenerational transmission of an early life health shock. Using extremely detailed Swedish data on the in utero disease environment of the first generation, and on various health and socioeconomic outcomes of the first and the second generations over a time span of 130 years, we seek to answer

three distinct questions. First, we estimate at what rate the impact of the initial health shock diminishes from one generation to the next. Second, we analyse whether there is a socioeconomic gradient in the intergenerational transmission of health. Third, we estimate the repercussions of the initial health insult on a variety of socioeconomic outcomes in the second generation. We contribute to the existing literature by shifting the perspective to the complete life course (as opposed to Almond et al., 2012) while exploiting less extreme health shocks than those caused by rare and devastating events (Richter and Robling, 2013, Van den Berg and Pinger, 2014). The first generation in our dataset was born in the decades around the turn of the 20th century – after the famine of 1866–8 (Doblhammer et al., 2011) and the last outbreak of smallpox in 1873–4 (Sköld, 1996), but before the Spanish flu pandemic in 1918 (Karlsson et al., 2014) – in an era characterised by gradual improvements in public health and by the absence of severe mortality crises.

Our analysis makes use of a unique and purpose-built Swedish dataset. A general challenge for an empirical analysis of this kind, even in Scandinavia, is that there are hardly any datasets which cover the long time spans and variables required without serious attrition due to mortality and migration.[1] We thus mainly rely on tailor-made datasets and match them to administrative data whenever possible. Our main dataset is a representative sample of 25,000 births from the cohorts 1930–34. It includes a large number of indicators for the parent generation – e.g. date and place of birth, socioeconomic and marital status, number of children – as well as a wide range of outcomes for the second generation – e.g. sex, place of birth, mortality, and socioeconomic outcomes in adulthood. All data are taken from official registers and the mortality database has exceptionally low attrition since it also includes mortality information on emigrants.

Our indicator on the first generation disease environment is the local infant mortality rate in the parish of birth of the mother. For 1,856 different birth parishes, we compiled information on annual IMR for the period 1880-1917. This variable is commonly used as a proxy for disease environment in utero and early childhood, and has been found to be an important predictor of adult health (Akachi and Canning, 2007, Bozzoli et al., 2009, Crimmins and Finch, 2006) and various other outcomes (Case and Paxson, 2009, Lawson and Spears, 2014). Some authors argue that post-neonatal mortality (PNM) is a better indicator of early life disease environment since it does not include neonatal mortality, which is also strongly associated with the access to pre- and perinatal care (Schmidt et al.,

[1] For example, the Swedish multigeneration register contains only individuals born after 1932 who were alive and registered at some point after 1960. Information on parents is complete only from the 1950 cohort onwards (Statistics Sweden, 2005). Moreover, existing demographic intergenerational databases (in Sweden, Canada, Italy, Switzerland and other countries) generally cover complete regions (e.g. a geographical cluster of parishes or a city) but are not representative of the entire population since the selection of areas do generally not take such criteria into account (Edvinsson, 2000). A main obstacle with existing demographic databases is also the lack of digitised individual level data for the period 1900-1950 (Bengtsson and van Poppel, 2011).

1995). PNM rates are not available for the time period we consider, but the drawback is probably less of a concern when considering conditions in Sweden in the late 19th and early 20th centuries, since access to those services was limited (Bhalotra et al., 2015) and their efficacy in improving infant health has been challenged (Pettersson-Lidbom, 2014). In auxiliary regressions, we confirm empirically that infectious disease is a much more important predictor of the IMR than access to health care. Nevertheless, we include a large battery of local fixed effects and trends in order to reduce the influence of persistence in local differences in public services and other confounding factors. Consequently the identifying variation we exploit are deviations of the IMR from local trends and levels.

A more serious concern regarding the use of the local infant mortality rate as an indicator of disease environment is the issue of selection (Almond et al., 2012). In particular if identification is driven by large deviations from the local mean, one should be concerned that the resulting sample will be very strongly selected (cf. Doblhammer et al., 2011). Any impact of local IMR on later-life outcomes would then be a combination of scarring and selection effects operating in opposite directions. Using a setting similar to ours, Hatton (2011) finds no evidence of selection, and Bozzoli et al. (2009) conclude that it may be more of an issue in developing than in developed countries. For the cohorts we study, the national IMR dropped from 10 per cent to 6 per cent during the observation period. This would be high by today's standards in developing countries, but much lower than the levels that have been experienced in developing countries in the past. However, it is very clear in our case that scarring dominates selection: the mothers who suffered an unfavourable disease environment had elevated mortality at older ages, and their children experienced worse health and SES outcomes.

In general, our results corroborate earlier studies on the importance of the in utero environment: we find that a maternal health shock significantly affects survival prospects in the second generation, that this effect is small or even non-existent during the first decades of life, but that it has a clear and significant impact on survival prospects after the age of 50 (Almond and Currie, 2011). The effect is also particularly pronounced for males: if the local IMR in the place of birth of the mother increases by ten percentage points, the risk of dying before the age of 70 increases by three percentage points (ten per cent) within this group. We do not find any evidence of this effect being driven by fertility responses.

In two respects our results however differ strongly from the typical findings of the previous literature. First, we find evidence of an inverse SES gradient in the intergenerational transmission of health. When the survival disadvantage becomes visible in the second generation from age 50 onwards, it is strongly concentrated among individuals with a better-than-average SES background. In particular males from a privileged background seem to be affected. This is in stark contrast with the previous literature, which almost

always finds that the intergenerational transmission is stronger in disadvantaged groups (Bhalotra and Rawlings, 2013, Costa-Font and Gil, 2013, Currie and Moretti, 2007, Kim et al., 2014).[2] By considering grandparental SES we are able to rule out the possibility that our results are driven by selection into the advantaged group. Instead, we provide evidence suggesting that this unusual result is attributable to environmental and behavioural factors in childhood and adulthood: looking at specific death causes, related in particular to cardiovascular diseases and diabetes, there is compelling evidence that in utero metabolic adaptations have taken place. Such changes are compatible with the thrifty phenotype hypothesis (Hales and Barker, 1992) and may be particularly likely to lead to metabolic disorders in an affluent environment (Barker, 1997). This candidate explanation is corroborated by surveys of dietary habits in Sweden in the years following the births of our second generation (Boalt, 1939, Odin, 1934). Besides, we find evidence suggesting that some of the gradient is driven by behavioural changes in adulthood.

Second, our results regarding the impact on labour market outcomes are equally intriguing. For earnings, we find a very strong impact of the maternal disease environment: a one-point increase in maternal IMR is associated with a one-percent reduction in adult earnings. The effect size is large: it is comparable to the impact found within the first generation by Lawson and Spears (2014) or to the effect of a ten-percent increase in birth weight (Black et al., 2007). Surprisingly, the effect is entirely driven by females. This result contrasts the findings in some previous literature which suggests that health shocks disproportionately affect the labour market outcomes for males (Black et al., 2007, Cai, 2010, Pelkowski and Berger, 2004). In our case, men's labour market outcomes are hardly affected, but elasticity of female earnings is -2 – a one percentage point increase in the maternal IMR is associated with a 2 per cent reduction of second generation female earnings in the 1970's. The effect appears to be particularly strong for females from low-SES backgrounds. We are able to show that this effect is linked to the expansion of the welfare state: for the relevant cohorts, female employment increased in particular in public services, and there appears to have been positive health-related selection into these professions.

The disadvantage associated with a maternal health shock is thus complex and multifaceted. But at what age is the disadvantage determined? Analysing educational attainment in the second generation, we find weak evidence that the disadvantage is fixed early in life: the maternal disease environment associates with shorter completed education for high-SES males and low-SES females – i.e. the groups that appear to be particularly affected – but the coefficients are generally small and not statistically significant. On the

[2] Almond et al. (2012) do find a similarly inverted black-white gradient, which they attribute to selection effects dominating amongst black mothers. This explanation is unlikely to apply in our case, since we find evidence of scarring in the first generation also for the disadvantaged group.

other hand, our analysis by death cause delivers relatively strong evidence of an epigenetic transmission. The recent epidemiological literature points to a relationship between "viral infections and preterm labour, and fetal congenital anomalies of the central nervous system and the cardiovascular system"(Mor and Cardenas, 2010).[3] According to the epigenetics story, environmental conditions in utero and early life may alter the genetic phenotype: for instance, if the disease load is high, bacteria could cross the placenta causing an inflammatory response syndrome which can have long-term consequences. There is some evidence that such phenotypic changes can also be transmitted across generations Hochberg et al. (2011).

The cohorts we study are special in the sense that they were born around the Great Depression, which evidently represents an environmental shock to the second generation. However, we do not find any evidence of the Great Depression having a detrimental effect on health and labour market outcomes of the cohorts exposed to it in early childhood. Using administrative data of the crisis impact at the local level, our difference-in-difference estimates for adult incomes are insignificant throughout. As regards health, we find evidence that the crisis had a protective effect on mortality from age 50 onwards. This finding is, on the one hand, consistent with the common result that downturns are associated with improvements in health (Gerdtham and Ruhm, 2006, Neumayer, 2004, Ruhm, 2000, 2003, 2004) but on the other hand deviates from the literature focusing on the early life enviroment, which generally finds adverse effects (Lindeboom et al., 2010, Van den Berg et al., 2009, 2011, 2006). We leave the investigation of this result as a topic for later research, but conclude that the effects of the crisis seem to apply across the board since we find no evidence that the crisis moderates the effects of the maternal health shock.

2 Data

One of the reasons why estimates of the intergenerational transmission of health and of the intergenerational effect of health on socio-economic outcomes are scarce is that data pose a significant challenge. Such an analysis requires tracking a sufficiently large number of individuals over the life-course and at the same time have information on the birth location and early-life disease environment of the parental generation. Other general challenges are attrition, mortality in early life and the traceability of migrants. The following two sections describe how we construct a tailor-made dataset which adresses most of these concerns.

[3]But also maternal stress has found to be a driver of preterm labour and, thus, of infant mortality (Goldenberg et al., 2008).

2.1 Individual Level Data

The base component of our dataset is a representative sample of individuals born in Sweden from January 1, 1930 to December 31, 1934, whom we can follow over the life course. We construct the micro-level data by digitising parish records from 133 parishes throughout the country on all individual births, including detailed information on birth date, place of birth, name, sex, mother's marital status, parents' name, date of birth and occupation.[4]

The Swedish church law of 1686 states that the clergyman in each parish should keep record for all children born in and out of wedlock. Vital statistics were thus introduced very early in Sweden and the information provided by parish records are generally seen as being of very high quality (Edvinsson, 2000). Covering everyone born during the sample period in the selected parishes the data includes 25,064 individual births corresponding to nearly 6 per cent of all births in Sweden in 1930-1934. The data also allow us to identify siblings born within this time period and twin births. There are 12,015 siblings in the dataset, from 5,279 mothers.

Figure 1 shows the geographical location of the birth parishes of our individuals. Birth parishes are distributed across all parts of the country and using information from the 1930 census we confirm that the locations covered in the dataset are representative of Sweden as a whole in terms of observable characteristics such as economic structure, average income and infant mortality, corroborating the external validity of our results.

Insert Figure 1 about here.

We follow each individual from birth to death, or to age 79-83 if still alive in 2013. Using name, gender and date and place of birth we trace the date of death of deceased individuals from the Swedish deathbook (*Sveriges dödsbok*). The deathbook contains the universe of deaths occurring in Sweden and to Swedes abroad during the period 1901-2013. In order to validate our matches we use information from cemetery burial records (Federation of Swedish Genealogical Societies, 2009).[5] As a second validation source we use tax records on annual incomes (labor, capital and business income) 2001–12 and the 1970 census.[6]

[4] Parishes (*församling/socken*) are subdivisions within the Church of Sweden. In 1930s there were about 2,200 parishes in the country.

[5] This source covers all grave registers in the country and includes more than 5 million entries.

[6] These sources also allow us to cross-check names, of particular importance for the correct tracing of women who often change their last name when getting married.

With the above sources we track 96 per cent of all individuals for a period of 79 years.[7] Among our cohorts 44.7 per cent did not survive to the age of 79. For deaths occurring after 1960 we also have information on individual death causes from the Death cause register (The National Board of Health & Welfare, 2014) covering all deceased residents who died in Sweden or abroad. Looking into descriptive statistics the quality of our death records seems very high. 2.6 per cent of all births in the sample are stillbirths, the infant mortality rate is 58 and child mortality 73 per 1,000 live births. This is all to be considered normal of the 1930s (Edvinsson et al., 2008) and suggests accurate and high quality recording. Similarly the sex ratio at birth is exactly the expected 1.05. Also with respect to death cause statistics we can confirm the external validity of our dataset: 38.5 and 30.6 per cent respectively of our sample population died because of circulatory diseases and cancer in the 1990s which corresponds to official figures for the corresponding demographic groups in the entire population (Statistics Sweden, 1997).

Our health outcome variables are a mutually exclusive set of binary mortality indicators over the whole life course (censored at age 79): mortality between ages 0–1; 1–50; 50–70; and 70 and older. Even though we consider individual-level mortality from birth onwards, we put particular emphasis on mortality between 50 and 70, as these ages have been identified as a critical period in the fetal origins literature.[8]

We also examine the intergenerational transmission of health to socio-economic outcomes.[9] Our main socio-economic outcomes are measured in adult life (age 36-40) when the second-generation individuals are typically active on the labour market after having completed their education. The primary socioeconomic outcome we consider is log individual earnings from the 1970 Census.[10] We also examine other labour market outcomes for the second generation, such as employment (total, full-time, part-time) and sector of employment. Finally, we examine the impact on completed years of schooling.

Table 1 provides descriptive statistics for the variables in the dataset.

[Insert Table 1 about here]

[7] The remaining 4 per cent are likely to be surviving emigrants; we thus include them as survivors in our analyses of mortality. Results are insensitive to the exclusion of this group.

[8] Barker's seminal work explores in particular the association between low birth weight and the incidence of certain types of diseases in middle age - among others, coronary heart disease, hypertension and diabetes (Barker, 1990, Hales and Barker, 1992).

[9] As we observe all individuals and either their income or their death (or both), we can examine the role of attrition. In particular we compare descriptive statistics of the sub-sample of individuals who died before 1970 to the ones who did not die before 1970 and we compare our main regression results with the corresponding specification estimated using IPW (available upon request). Importantly, attrition does not seem to be an issue in our data.

[10] As discussed by e.g. Haider and Solon (2006) and Böhlmark and Lindquist (2006), income data measured at ages 30-45 generally provide a good proxy of lifetime income as they are less likely to fluctuate due to life-cycle biases.

The parish records provide pertinent information on the parents (marital status, name, date of birth and occupation). As household socio-economic status may play an important role for health and socioeconomic outcomes of the second generation, we classify the parents' occupations according to the HISCO system (Historical International Standard Classification of Occupation) to derive relevant SES groups. The HISCO rank occupations according to the required skill level of an occupation, where 0 indicates the most skilled non-manual jobs, and 9 indicates the lowest skilled manual occupations. Table 2 lists the major HISCO groups, together with a short description and the absolute and relative frequencies in our data. Many household heads belong to HISCO group 6, representing mainly farmers, and to HISCO group 9, representing workers employed in unskilled jobs.[11]

[Insert Table 2 about here]

To allow for a survival analysis of the first generation we also collect information on the date of death of the mothers from the death book. We identify the exact date of death of 89 per cent of our first generation population.

Digitising the parish records gives an advantage compared to official multigenerational registers since the latter only contain individuals who were alive and registered in Sweden at some point in time after 1960. Similarly information on the parental generation is conditioned on being alive and parental data is complete only from the 1950 cohort and onwards (Statistics Sweden, 2005). Using parish records also assure that we do not have any misreporting in the place of birth of an individual.[12] These unique features of the dataset allow us to match individual-level data for both the first and the second generation to detailed information on the local economic and in utero disease environment.

2.2 Maternal Disease Environment In Utero

Our indicator of the maternal disease environment in utero is the infant mortality rate in her parish of birth. Lifetime individual-level data have previously been combined with local IMR to evaluate the impact of exogenous variation in early-life conditions of one

[11] The reported information is generally based on fathers' occupation.

[12] In official registers the parish of birth reported for cohorts born until 1946 refers to the place of the actual birth of an individual, i.e. if an individual was born in a hospital the parish of birth reported refers to the location of the hospital in which someone was born (Skatteverket, 2015). The transition to institutional delivery started in the late 1920s and was initially very smooth, but in the mid-1940s the majority of births took place out of the home (Wisselgren, 2005)

generation (see e.g. Bengtsson and Lindström, 2003, Van den Berg et al., 2009).[13] To analyse the role of in utero disease environment of a parental generation we first use the information provided on parents in the parish birth records of the second generation and the death book to exactly identify the parish of birth of the mothers of the children of the second generation. The first generation mothers are born between 1880 and 1918 in 1,864 parishes across the country.

We calculate the IMR within the time window 1880-1918 on the parish level. For the period 1880-1900 we use IMR data from *Statistiska Centralbyrån* (Statistics Sweden, SCB). For the remaining years we calculate local IMR from the deathbook.[14] We obtain the IMR at birth location for 99.8 per cent of the mothers in the first generation.

Figure 2 shows the average regional variation in the IMR level from 1880 to 1918, while Figure 3 shows the change over time of the IMR from 1880 to 1918 averaged over all parishes.[15]

Insert Figure 2 about here.

Insert Figure 3 about here.

2.3 Economic Environment

The cohorts of 1930-1934 were born around the Great Depression, which represents an additional environmental shock. The peak of the crisis in Sweden appeared in March 1932 with low per capita incomes and high unemployment (Kobrak and Mira, 2013), but the economy recovered remarkably fast: in 1934, the gross domestic product was already back to its 1930 level (Mitchell, 1998). To test if the recession had any effects on health and socioeconomic outcomes we add local level information on the general economic situation measured as deviations in annual municipality income tax revenue for the period

[13] Along the same lines we use the infant mortality rate as a proxy for the health environment at birth and/or in utero. Although its effect might be theoretically ambiguous, as only surviving infants are considered, it accounts for huge differences in the cross-cohort health outcomes and it has been shown to be a good proxy for early health environment (mortality and height; see: Bozzoli et al., 2007, Crimmins and Finch, 2006). We further discuss the potential determinants of IMR in Section 4.

[14] The IMR for the parish of birth of the mother is missing for 16.9 per cent of the mothers in our sample. For these cases we impute IMR based on the regional annual IMR, taking the annual IMR in the region of birth and subtracting a weighted average of all available IMRs at the parish level in the same region. In addition, information on the parish of birth of the mother is missing for 14.6 percent of our first generation. In these cases we impute the IMR of the parish of birth of the firstborn child. Such imputed measures are cruder than the local IMR but improves representativeness and allows us to include additional observations.

[15] As described below we use fixed effects implying that the identifying variation in our estimation will correspond to deviations from the levels and time trends in IMR within each parish.

1930-1934, collected from yearbooks (Statistics Sweden, 1935).[16] We divide the yearly tax revenues at the municipality level by the working population in 1930 (from the 1930 Census) and we deflate this measure by the Cost of Living Indicator (CLI, by Statistics Sweden). We define the crisis indicator $crisis^i_{sy}$ as the negative logarithmic deviation of the deflated per capita tax revenues (tax) in year $y = (1931, \ldots, 1934)$ with respect to the same measure in 1930, for each parish s.

$$crisis^i_{sy} = -\left[log(tax_{s,y}) - log(tax_{s,1930})\right], \quad y \in \{1931, \ldots, 1934\} \tag{1}$$

If the indicator is positive the crisis hit a parish, and if negative the economic situation in a parish was better in year y than in 1930. Figure 4 shows how the indicator changes over time and across parishes. On average, the indicator increases until 1932 and then declines again until 1934.

[Insert Figure 4 about here]

3 Empirical Strategy

As Figure 3 reveals, Sweden experienced a declining infant mortality rate over the relevant period. In addition to these temporal trends, there are also spatial patterns. If these patterns are not taken into account, the IMR coefficient simply reflects the fact that mothers who were exposed to higher infant mortality rates are older (confounding time-trend) and/or were born in places where, for example, public services were scarce or had a low quality or, again, that individuals select themselves into municipalities according to their health endowment. It could also be that different regions in Sweden exhibit different rates of technological progress or expansions of the medical sector.

We control for regional fixed effects and different regional-specific trends (spatial time trends and ageing trends in the first generation) using the following baseline econometric model, similar to the one employed by Almond et al. (2012):

$$Y^{ij}_{ptrsya} = \alpha + IMR^j_{pt}\beta_1 + X^i\gamma + \theta^j_a + \theta^j_a t^j + \delta^j_r + \delta^j_r t^j + \epsilon_{ij} \tag{2}$$

[16] At the time parishes and municipalities more or less coincided. The 133 parishes covered by our individual-level dataset are grouped in 118 municipalities.

where Y^{ij}_{ptrsya} refers to health or to socioeconomic outcomes for individual i, born in parish s in year y, belonging to the second generation. j indicates individual i's mother, belonging to the first generation, while p refers to mother j's parish of birth, t her year of birth, r the region of birth and a age when giving birth. IMR^{j}_{pt} proxies the disease environment that the individuals in the first generation experienced in utero. We use a linear probability model (LPM) to estimate Eq. 2 when the focus is on binary mortality indicators. For our economic indicators we estimate Eq.2 by OLS.

The coefficient β_1 reports the effect of the health shock in the first generation, IMR^{j}_{pt}, that is our effect of interest. We argue that this coefficient is identified conditional on the covariates.[17]

X^i is a vector of control variables pertaining to individual i in the second generation, born in parish s in year y. It contains binary indicators for sex, whether a twin birth occurred or whether individual i is born in wedlock and dummy series for the order of birth, the quarter and year of birth, the parish of birth of individual i and the occupation of the household head. We also control for a crisis indicator to capture possible effects of the Great Depression on mortality. Furthermore we control for spatial effects on the parish level where the second generation is born and for seasonal variation.

Equation 2 also accounts for several characteristics of the first generation. We include maternal age fixed effects θ^{j}_{a} to account for the influence of the mother's age on the foetus' health and mother's region of birth r (*Län*) δ^{j}_{r} fixed effects. We control for time trends using two vectors. The first ($\theta^{j}_{a}t^{j}$) addresses time trends in the age of the mother. It takes into account that the age effect of mothers at birth is changing over time. The second time trend ($\delta^{j}_{r}t^{j}$) accounts for separate time trends by region. Holding trends and levels within regions fixed, β_1 is thus identified by deviations from those trends and levels.

Another important issue is the population for whom the effect is identified. Even if the health shock in the first generation is random, mothers who are hit by the most severe shocks might die younger or might not get an offspring. We can only identify the effect for those first-generation individuals who survived and become fertile. Due to selection occurring among the first generation our estimate is likely a lower bound. In Section 4.4 we elaborate further on the transmission channels in order to be able to better assess the quality of the regression results.

[17] In Section 4 we discuss whether IMR is a good approximation for the health environment of the mother at birth and whether it is an appropriate measure for the effects that we are considering in the second generation.

4 Results

This section presents the results from the empirical analysis. We first provide some descriptive and graphical evidence for the main explanatory variable and the main outcomes. We then turn to regression estimates for the second generation. Finally, we deal with issues related to selection and confounding factors, and investigate the impact of the original health shock within the first generation.

4.1 Explorative Graphical Analysis

We now provide some visual evidence regarding the relationship between the maternal disease environment and survival prospects in the second generation. Since infant survival has been shown to be affected by intergenerational transmission (Almond et al., 2012), we start by zooming in on the first 365 days of life. Figure 5 provides two hazard plots. The left-hand side figure shows the cumulative hazard (by 15-day bins) for children born to mothers from different health environments (defined in this case as a dummy variable indicating positive and negative deviations from the local trend), whereas the right-hand side shows the difference in the cumulative hazards between the two environments.

[Insert Figure 5 about here]

The figure does not deliver any evidence supporting the hypothesis that children born to disadvantaged mothers fare worse in terms of early life survival. We can bound the overall effect on second-generation IMR at less than 0.5 percentage points in any direction, and the point estimate is everywhere very close to zero.

Figure 6 shows the same exercise for the entire life course. One again the evidence suggests that the impact of the maternal disease environment is close to zero for a long time – but from age 50 onwards there is evidence of an increasing penalty for children born to mothers from a poor disease environment.

[Insert Figure 6 about here]

4.2 Second Generation Mortality

The descriptive evidence from the above subsection suggests that there is no manifestation of an impact of a disadvantageous maternal disease environment during the first

years of life, whereas there appears to be a growing disadvantage starting around the age of 50. This evidence is consistent with the Barker hypothesis, which suggests that an adverse fetal programming may cause various health problems from middle age onwards. We now formally test this in a regression framework while controlling for various environmental factors which could possibly have confounded the relationship between the maternal disease environment and second-generation outcomes. Table 3 presents the results.

[Insert Table 3 about here]

Each specification presents results for second-generation mortality during a particular period in life, conditional on survival up to that period. In addition to the main explanatory variable – the mother's disease environment (IMR^j) – we also control for the child's sex, the degree of crisis in the birth parish in the birth year, and a set of fixed effects for the birth parish, the mother's birth region, and the mother's birth year. Clearly, only the mortality rate between 50 and 70 is significantly correlated with the mother's disease environment. According to this estimate, each point increase in the IMR of the maternal birth parish associates with an 0.14-point increase in the mortality rate. Since the standard deviation of the maternal IMR is 0.05, a one standard deviation increase in this variable is associated with an increase in the second generation mortality by 0.7 percentage points. With a baseline risk is 16.4 per cent, this effect is relatively large.

We take two main messages out of Table 3. First, there is limited evidence of selective mortality before age 50: the relatively small and insignificant effects in the first two columns cancel out to some extent. Second, there is a strong impact of the maternal disease environment on mortality between 50 and 70. With a p value of 0.023, this result would remain statistically significant at conventional levels also after a correction for multiple testing (p value of 0.092 using the Bonferroni method).

The existing literature suggests that males are more sensitive to health insults than females, and that the socioeconomic environment into which the child is born can act as a buffer to mitigate some health shocks. In Table 4, we allow for a wide range of robustness checks and analyses by subpopulation.

Each column of Table 4 adds some additional control variables in order to assess the robustness. The rows denoted A-I represent different subsamples, defined by sex and the socioeconomic status of the head of the household in the first generation.[18] Clearly, the survival disadvantage observed in Table 3 is robust to the inclusion of region-specific trends in the first generation, parental occupation, and birth order effects. The parameter drops slightly but remains statistically significant at the 5 per cent level.

[18] We will return to the potential endogeneity of this variable in section 4.4.

[Insert Table 4 about here]

We also find evidence of effect hereogeneity. When splitting the sample by sex, it becomes clear that the negative effect is entirely driven by males. When splitting by parental SES, the effect seems to be concentrated in higher socioeconomic groups. When we interact the two dimensions, the effect is particularly pronounced amongst males born in families of relatively high SES. An increase in the first-generation IMR by one standard deviation increases mortality between 50 and 70 by almost two percentage points (1.91) in this group. Measured against a baseline rate of 18.5 per cent, this effect is indeed sizeable.[19]

What are the mechanisms responsible for these findings; in particular the inverted SES gradient amongst males? We start by considering specific death causes, and then turn to socioeconomic outcomes. Since we find that the original health insult appears to interact with the early life environment of the second generation – as captured by the household head SES – it seems reasonable to consider leading death causes which are likely to have a genetic component while at the same time being modifiable. For instance, data from the death cause register enable us to differentiate between cardiovascular diseases with concomitant risk factors[20] and other cardiovascular diseases with no modifiable risk factors.

[Insert Table 5 about here]

Several things are immediately clear from the results in Table 5. First, cardiovascular (CDV) diseases without concomitant risk factors (Panel B) appear to be responsible for a big share of the intergenerational transmission. This death cause, responsible for 23.7 per cent of the total deaths between 50 and 70, explains about 50 per cent of the transmission, and the concentration amongst males and in the group with higher SES is very similar as for all-cause mortality. In addition, an inverted SES gradient concentrated amongst males is supported by the estimates for the category 'diseases of the endocrine and digestive system' which includes diabetes as the primary cause of death.

These findings are informative regarding the mechanism driving the results, since these health conditions – cardiovascular disease and diabetes – and the related risk factors hypertension and obesity are the main outcomes of adverse fetal programming according to the Barker hypothesis (Barker, 1990). Importantly, such metabolic adaptations in

[19] In the interest of clarity, we present effect heterogeneity results as split-sample regressions. The effect heterogeneity results we present are statistically significant. Results are also robust to the inclusion of fixed effects for the mother's birth parish: due to the large overlap with the second-generation birth parish, we left this variable out of the main specification.

[20] We have information about hypertension, diabetes and alcohol consumption. Conventional risk factors for cardiovascular diseases that can be modified and/or treated are: having a diabetes condition, hypertension condition, unhealthy dietary habits, alcohol consumption, smoking and physical inactivity (see e.g. Khot et al., 2003, for an overview). Unmodifiable risk factors include for instance age and family history.

the second generation are compatible with intergenerational transmission mechanisms (Langley-Evans, 2015).

The onset of metabolic problems is likely determined early in life and several epidemological studies show that a high-nutrient diet in infancy, in particular diets rich in dairy proteins, associates with programming of the metabolic syndrome in children – specifically with increased BMI and obesity in childhood (Pearce and Langley-Evans, 2013, Weber et al., 2014). A household survey conducted 1936–7 provides detailed information on dietary habits in Swedish families during the first years of life of the second generation. The survey concludes that young children from higher SES background were given on average more nutritious food than children from lower SES background. The survey also shows that milk and, more generally, dairy products represent the type of protein that is consumed the most. Also in this case there is clear socioeconomic difference among children: a five year old child in a middle- or high-SES household on average consumed 4.6 litres milk per week – 25 percent more compared to a peer from a low-SES household (Boalt, 1939).[21] Similar indications are provided by larger, but less detailed, surveys conducted in the 1930's (Medicinalstyrelsen, 1934, Socialstyrelsen, 1938). Also, while sugar never previously had been an everyday consumption good, it becomes all the more common among higher SES-housholds in Sweden during this time period (Bolin, 1934, Torell, 2013). It is thus possible to explain a substantial part of the inverted socioecomic gradient with reference to the *thrifty phenotype*: second generation high-SES individuals were more likely to be exposed to a calorie rich diet and a sedentary lifestyle than low-SES individuals during their childhood.

However, Table 5 suggests that also behavioural factors matter to the intergenerational transmission. For example cancers appear to be responsible for some of the male penalty in intergenerational transmission, and the most prevalent cancers affecting males have a strong behavioural component. Lung and oral cavity cancers appear to counteract the intergenerational transmission, and the estimated reduction in the prevalence of this death cause is large, considering that less than one per cent of the sample died between 50 and 70 due to this cause. Besides, the absence of an SES gradient in females appears to be attributable to some extent to a reduction in cardiovascular disease with concomitant risk factors within the high-SES group.

We hence conclude that the fetal programming to some extent also appears to interact with behavioural changes in adulthood. This is to be expected: the birth cohorts under study were some of the first to become aware of the perils of smoking in adult ages (the Surgeon General's report on smoking and health was published in 1964; Holford

[21] The survey was conducted by *Kooperativa Förbundet* with the aim to map dietary habits across socioeconomic groups, and covered 378 housholds (1163 individuals) across the countrey. Each houshold filed a protocol for each meal and every food intake for every household member during seven days.

et al., 2014). Likewise, these cohorts were the first to be exposed to large-scale prevention programmes for cardiovascular disease, which are believed to have contributed to a reduction in morbidity and mortality (Weinehall et al., 1999).

The inverse socioeconomic gradient in the intergenerational transmission may consequently be due to a combination of environmental and behavioural factors in childhood and adulthood being of greater importance for males. This line of reasoning seems to be supported by trends in mortality rates, as Figure 7 illustrates.

[Insert Figure 7 about here]

Figure 7 shows that males exhibit a strong downward trend in mortality as well as a strong socioeconomic gradient, whereas none of these features is visible for females. Our tentative conclusion is thus that the convergence of male mortality rates with female rates which this generation experienced, was to a large extent driven by improvements for higher-SES individuals who did not carry a disadvantage related to a health shock in the previous generation. This explanation is very similar in spirit to the ideas brought forward by Cutler et al. (2006), who postulate that new knowledge and treatment possibilities will generally benefit the higher SES groups first, so that the SES gradient may widen in some periods. Interestingly, the SES gradient in the intergenerational transmission of health may move in the opposite direction for exactly the same reasons.

4.3 Second Generation SES Outcomes

We now turn to regression estimates for socioeconomic outcomes of the second generation, measured at the suitable time when our individuals were between 36 and 40 years old (cf. Böhlmark and Lindquist, 2006). The SES outcomes we now consider may be seen either as mechanisms possibly explaining the findings we have reported for mortality above, or as outcomes in their own right.

Table 6 provides regression results for a range of specifications. Each column is equivalent to the corresponding columns in Table 4 and the split-sample estimates presented in the rows are also consistently defined. The number of observations differs slightly due to migration (those migrating out of Sweden before 1970 are not observed) and mortality (individuals who died between 1970 and the age of 50 are included here but not in Table 4).

[Insert Table 6 about here]

The statistical significance is generally somewhat weaker for earnings compared to mortality, but a clear picture nevertheless emerges: there is evidence of a relatively large penalty associated with the maternal health shock, and the result is mainly driven by females and individuals from lower socioeconomic groups. In particular females seem to suffer from the maternal health shock: the negative earnings impact for this group is more than twice as large as the overall impact. A standard deviation change in the maternal disease environment associates with a five-percent reduction in earnings in the overall population, and with a 10-per cent reduction for females. Interestingly, there is no evidence of a disadvantage for males within the high SES group. Thus, even though the maternal disease environment leads to elevated mortality later on for this group, there is no evidence that this disadvantage is manifested in earnings in middle age.

Considering labour market participation, Table 7 presents results for three employment variables – for the overall sample and for males and females separately. The findings for earnings are reflected in the results for employment and, in particular for the female subgroup, a poor maternal disease environment associates with a reduction in employment.

[Insert Table 7 about here]

The results for 1970 earnings thus produce an additional intriguing result – that the economic disadvantage of a poor maternal disease environment is disproportionately suffered by women – but no support for the previous finding for high-SES males. In an attempt to resolve the issue we turn to education. Table 8 reports estimates with years of schooling (measured in 1970) as the outcome variable. Estimates are not statistically significant at conventional levels and do not show a consistent pattern when looking at effect heterogeneity. In contrast to Richter and Robling (2013) we do not note any evidence of an education effect of the maternal health insult and we again find evidence suggesting that most of the effects of this kind of health insult manifest themselves only in adulthood.

[Insert Table 8 about here]

Since we can rule out that the maternal health insult had a large impact on choices regarding education early in life, the strong earnings disadvantage of low-SES females must have other causes. The cohorts we consider lived through the expansion of the welfare state and the associated improvements in labour market opportunities for females (Magnusson, 2010). It thus seems natural to hypothesise that this trend expanded the opportunities disproportionately for females, and that the transition into formal employment

within this group was related to health. In Table 9, we formally test this hypothesis by regressing the maternal disease environment on employment in public services. Estimation results are presented in Table 9.

[Insert Table 9 about here]

The results do indeed support a story according to which the emerging welfare state selectively employed females (in particular low-SES females) with a positive health transmission from the previous generation. We may thus conclude that the female penalty is driven by changes on the labour market, whereas the male penalty is largely unrelated to education and labour market performance.

4.4 Tracing the Health Insult

Since the variation exploited for the identification of an intergenerational transmission of health is non-experimental, it is important to consider omitted variables and other confounding factors which may bias our results. Besides, it is important to understand where the variation in the in utero maternal health environment comes from. Selective mortality in the second generation was addressed in section 4.1. However, we conditioned on variables which are determined only in adulthood for the first generation, and which may thus well be affected by the original health insult. This would lead to the familiar bad control problem (Angrist and Pischke, 2008) and thus it is important to investigate this issue further.

[Insert Table 10 about here]

In a first step, we analyse whether relevant second-generation observables at birth vary systematically with the maternal disease environment. Table 10 presents estimates that are all very small and generally insignificant. Consequently there seems to be no effect on first-generation SES, but given that we find strong effects on second-generation SES (Tables 6 and 7) we want to further pursue this issue. Reliable information on the grandparental SES is not available, but the mothers' maiden names are contained in the data, and these do to some extent signal social position (Clark, 2012). Similar to Clark, we extract surnames signalling higher classes – noble names; latinised surnames; typical bourgeois names – to construct a very crude measure of the SES of the maternal grandparents. The bulk of noble surnames dates back to the 17^{th} and 18^{th} centuries, whereas the latinised surnames are partly from the pre-industrial period and partly from the decades

around 1900. Bourgeois surnames typically date from the 19th century. Our categorisation thus likely captures a combination of the grandparental SES and the SES of earlier ancestors. Given the low degree of social mobility before the industrial era (Clark, 2012, Lundh, 1999), this might be an acceptable proxy for the social origin of the mother.

Figure 8 plots histograms of the maternal disease environment by grandparental SES. The health insult suffered by the mother is symmetric and almost equally distributed for those coming from low- and high-SES families, respectively.

[Insert Figure 8 about here]

Thus, the original health shock appears to be unrelated to SES. Given rigid social structures in the parental generation, one might also conjecture that the same holds for the parental SES, even though it is determined after the health insult. We investigate this hypothesis in Figure 9 by contrasting socially mobile individuals to the rest. Again, we do not find any evidence that the maternal disease environment is related to SES in either of the two generations.[22]

[Insert Figure 9 about here]

Next, we investigate the *nature* of the health shock. Relying on data from the 19th and early 20th century is a clear limitation since the available information on local public health conditions is very limited. This may however also be seen as an advantage in the sense that the local variation over time was much more random in those days, when health care services were typically neither available nor effective, and a clear socioeconomic gradient in health had not yet emerged (Bengtsson and Dribe, 2011). We now conduct an analysis at the regional level ($N = 25$) for the time period 1890-1910 and regress the regional infant mortality rate on various potential determinants, including infectious diseases and the proportion of children vaccinated against smallpox.[23]

[Insert Table 11 about here]

The estimates in Table 11 suggest that infectious disease is the main driver of variation in regional IMRs, whereas the availability of health care resources appear to be of less

[22] We also reran all regressions above using grandparental SES instead of parental SES and even though some precision is lost due to the lower informativeness of the grandparental SES indicator, all results were qualitatively the same.

[23] The last outbreak of smallpox in Sweden occurred in 1873–4, i.e. some years before the oldest mothers were born (Sköld, 1996). Thus, the vaccination variable serves as a proxy for the local health infrastructure and not for the disease environment.

importance. We also control for a proxy of living standards using the log of farm workers' wages (cf. Lundh and Prado, 2015), but it does not seem to be a relevant driver of regional IMR.

We further try to shed some light on whether our estimates capture maternal exposure in utero or during the first year of life. So far, all regressions have used the local IMR in the year of birth of the mother – and thus possibly a combination of in utero and neonatal exposure. For mothers born after 1900 we have more exact information, which allows us to exactly time the impact of the local IMR exposure. Table 12 compares the overall results to the results of the subset of mothers with exact date IMR information. The results clearly indicate that a large share of the estimated IMR effect is related to in utero exposure.

[Insert Table 12 about here]

In a last set of specifications we analyse the impact of the maternal disease environment in the first generation. Table 13 provides the results for our main mortality outcome (mortality between ages 50 and 70). We use the indicator for grandparental SES specified above to assess whether there is heterogeneity in the impact. The overall estimated impact is small in magnitude (a one-SD increase in IMR associates with a *decrease* in mortality by 0.02 percentage points) but the estimates suggest there is some heterogeneity with respect to social class: mothers of higher SES appear to suffer an increase in mortality when exposed to an unfavourable disease environment in utero, while the opposite is true for mothers from lower SES. However, the estimates are very imprecise and thus no definite conclusion may be drawn, even though the evidence seems to suggest that the impact is weaker in the first generation than in the second generation.

[Insert Table 13 about here]

Thus, there is no evidence that ages 50-70 represent a particularly critical period in the first generation. But this does not necessarily imply there is no effect in adulthood at all. In order to formally test this, we conduct survival analysis for the mortality hazard of mothers from entering the sample onwards. Table 14 shows the results. The findings in this part clearly show evidence of scarring dominating selection in the first generation: mothers exposed to a negative disease environment did suffer elevated mortality rates in adulthood and this effect is driven by the lower socioeconomic groups. Results from Table 12 and Table 13 suggest that a socio-economic gradient in mortality appears already in the first generation – but going in the opposite direction compared to the second generation.

[Insert Table 14 about here]

5 Conclusion

The issue of persistence of disadvantage within families has attracted great interest in economic research (Hertz et al., 2007, Lindahl et al., 2012, Mazumder, 2005). Still, evidence on the intergenerational effects of health remains scarce. More knowledge on this topic seems urgent, not the least since investments in maternal health potentially could have large returns that accumulate across generations.

This paper contributes to a small literature on the intergenerational transmission of health and its gradient. Using historical data from Sweden for individuals born between 1930 and 1934 and their parents, and exploiting a natural health shock predicting the in utero health environment of the first generation – deviations from the local infant mortality rate – we examine the intergenerational transmission of in utero health on second generation health and socioeconomic outcomes.

In accordance with the fetal origins hypothesis the results suggest that health shocks hiting the first generation more than 100 years ago are still present in the second generation's health outcomes and thus still shape today's society: A one-standard deviation change in the mother's health environment causes the hazard to die between age 50 and 70 to increase by 0.7%.

While our results are mainly driven by males (in line with the existing literature), our data reveal an intriguing inverted SES gradient. Analysing death cause data, we find corroborating evidence that the effects to a large extent are driven by cardiovascular diseases and other conditions which have been linked to the in utero environment. The inverted SES gradient may thus be indicative of adverse effects from a rich diet (particularily during childhood), suffered by individuals exposed to poor conditions in utero. At the same time these data also suggest that the intergenerational transmission of health might be influenced by a behavioural component.

Examining socio-economic outcomes we do not find additional evidence of an inverted socio-economic gradient. On the contrary, evidence points to that females and individuals from lower SES suffer an earnings disadvantage due to the intergenerational health transmission (a one-SD change in the IMR in the first generation associates with a 10-per cent reduction in earnings for females). Females from low SES backgrounds suffer from particularly unfavourable labor market outcomes: they are less likely to be employed and those who suffer a health insult from the previous generation are also less likely to be employed in the public sector. However, analysing years of education as an outcome we conclude that the detrimental outcomes in the second generation do not seem to emerge before individuals enter the labour market. The female penalty thus seems driven by

the labour market, while the male penalty is largely unrelated to education and labour market participation.

Lastly we examine the impact of the health shock on the first generation: while there is no one-to-one transmission of the adverse effect on mortality between ages 50 and 70, when looking at the overall mortality distribution, there is some evidence that a worsening of the disease environment at birth increases the mortality risk for first generation mothers from a low SES family.

All in all, our investigation provides evidence that intrauterine programming is not only confined to one generation. It is inherited non-genetically from mothers to their children. In contrast to the existing literature that focuses on early life mortality outcomes, this paper shows that the effect on later life mortality might be even more relevant.

References

Aizer, A. and Currie, J. (2014). The intergenerational transmission of inequality: Maternal disadvantage and health at birth. *Science*, 344(6186):856–861.

Akachi, Y. and Canning, D. (2007). The height of women in Sub-Saharan Africa: The role of health, nutrition, and income in childhood. *Annals of Human Biology*, 34(4):397–410.

Almond, D. (2006). Is the 1918 influenza pandemic over? *Journal of Political Economy*, 114(4):672–712.

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

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

Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124(4):1729–1772.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

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

Barker, D. J. (1997). Maternal nutrition, fetal nutrition, and disease in later life. *Nutrition*, 13(9):807–813.

Bengtsson, T. and Dribe, M. (2011). The late emergence of socioeconomic mortality differentials: A micro-level study of adult mortality in southern Sweden 1815–1968. *Explorations in Economic History*, 48(3):389–400.

Bengtsson, T. and Lindström, M. (2003). Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766–1894. *International Journal of Epidemiology*, 32(2):286–294.

Bengtsson, T. and van Poppel, F. (2011). Socioeconomic inequalities in death from past to present: An introduction. *Explorations in Economic History*, 48(3):343–356.

Bhalotra, S., Karlsson, M., Nilsson, T., et al. (2015). Infant health and longevity: Evidence from a historical trial in Sweden. Technical report, Institute for Social and Economic Research.

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

Bhalotra, S. R. and Venkataramani, A. (2011). The captain of the men of death and his shadow: Long-run impacts of early life pneumonia exposure. *Available at SSRN*.

Bharadwaj, P., Løken, K., and Neilson, C. (2011). Early life health interventions and academic achievement. *American Economic Review*.

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

Boalt, C. (1939). *27.000 måltider: en undersökning av kostvanor*. Kooperativa förbundet.

Böhlmark, A. and Lindquist, M. J. (2006). Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *Journal of Labor Economics*, 24(4):879–896.

Bolin, I. (1934). *Sockret. Vårt billigaste födoämne*. Gothenburg, Sweden, Gø"teborgs litografi AB.

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

Bozzoli, C., Deaton, A. S., and Quintana-Domeque, C. (2007). Child mortality, income and adult height. Working Paper 12966, National Bureau of Economic Research.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, 17(1):77 – 90.

Case, A. and Paxson, C. (2009). Early life health and cognitive function in old age. *The American Economic Review*, 99(2):104.

Clark, G. (2012). What is the true rate of social mobility in Sweden? A surname analysis, 1700-2012. *Manuscript, Univ. California, Davis*.

Classen, T. J. (2010). Measures of the intergenerational transmission of body mass index between mothers and their children in the United States, 1981–2004. *Economics & Human Biology*, 8(1):30–43.

Costa-Font, J. and Gil, J. (2013). Intergenerational and socioeconomic gradients of child obesity. *Social Science & Medicine*, (93):29–37.

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

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

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives*, 20(3):97–120.

Doblhammer, G., van den Berg, G. J., and Lumey, L. H. (2011). Long–term effects of famine on life expectancy: A re–analysis of the great Finnish famine of 1866–1868. *IZA Discussion Papers*, 5534.

Edvinsson, S. (2000). The demographic data base at Umeå university: A resource for historical studies. *Handbook of International Historical Microdata for Population Research*, pages 231–248.

Edvinsson, S., Gardharsdóttir, Ö., and Thorvaldsen, G. (2008). Infant mortality in the Nordic countries, 1780–1930. *Continuity and Change*, 23(03):457–485.

Federation of Swedish Genealogical Societies (2009). Buried in Sweden (Begravda I Sverige).

Gerdtham, U.-G. and Ruhm, C. J. (2006). Deaths rise in good economic times: Evidence from the OECD. *Economics & Human Biology*, 4(3):298–316.

Goldenberg, R. L., Culhane, J. F., Iams, J. D., and Romero, R. (2008). Epidemiology and causes of preterm birth. *The Lancet*, 371(9606):75–84.

Haider, S. and Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, 96(4):1308–1320.

Hales, C. N. and Barker, D. J. P. (1992). Type 2 (non-insulin-dependent) diabetes mellitus: The thrifty phenotype hypothesis. *Diabetologia*, 35(7):595–601.

Hatton, T. J. (2011). Infant mortality and the health of survivors: Britain, 1910–50. *The Economic History Review*, 64(3):951–972.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2007). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, 7(2).

Hochberg, Z., Feil, R., Constancia, M., Fraga, M., Junien, C., Carel, J., Boileau, P., Le Bouc, Y., Deal, C., Lillycrop, K., Scharfmann, R., Sheppard, A., Skinner, M., Szyf, M., Waterland, R., Waxman, D., Whitelaw, E., Ong, K., and Albertsson-Wikland, K. (2011). Child health, developmental plasticity, and epigenetic programming. *Endocr Rev*, 32(2):159–224.

Holford, T. R., Meza, R., Warner, K. E., Meernik, C., Jeon, J., Moolgavkar, S. H., and Levy, D. T. (2014). Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964-2012. *Jama*, 311(2):164–171.

Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., and Eriksson, T. (2006). American exceptionalism in a new light: A comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. *IZA discussion paper*.

Karlsson, M., Nilsson, T., and Pichler, S. (2014). The impact of the 1918 spanish flu epidemic on economic performance in sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of health economics*, 36:1–19.

Khot, U. N., Khot, M. B., Bajzer, C. T., Sapp, S. K., Ohman, E. M., Brener, S. J., Ellis, S. G., Lincoff, A. M., and Topol, E. J. (2003). Prevalence of conventional risk factors in patients with coronary heart disease. *Jama*, 290(7):898–904.

Kim, Y., Sikoki, B., Strauss, J., and Witoelar, F. (2014). Intergenerational correlations of health among older adults: Empirical evidence from Indonesia. *The Journal of the Economics of Ageing*, (0):–.

Kobrak, C. and Mira, W. e., editors (2013). *History and Financial Crises: Lessons from the 20th Century.* Routledge, New York.

Langley-Evans, S. (2015). Nutrition in early life and the programming of adult disease: A review. *Journal of Human Nutrition and Dietetics*, 28:1–14.

Lawson, N. and Spears, D. (2014). What doesn't kill you makes you poorer: Adult wages and the early-life disease environment in India. *World Bank Policy Research Working Paper*, (7121).

Lindahl, M., Palme, M., Sandgren Massih, S., and Sjögren, A. (2012). The intergenerational persistence of human capital: an empirical analysis of four generations.

Lindeboom, M., Portrait, F., and Van den Berg, G. J. (2010). Long-run effects on longevity of a nutritional shock early in life: The dutch potato famine of 1846–1847. *Journal of health economics*, 29(5):617–629.

Lundh, C. (1999). The social mobility of servants in rural Sweden, 1740–1894. *Continuity and Change*, 14(01):57–89.

Lundh, C. and Prado, S. (2015). Markets and politics: The Swedish urban–rural wage gap, 1865–1985. *European Review of Economic History*, 19(1):67–87.

Magnusson, L. (2010). Sveriges ekonomiska historia.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *Review of Economics and Statistics*, 87(2):235–255.

Medicinalstyrelsen (1934). *Socialhygienisk undersökning i Västerbottens och Norrbottens län 1929-1931*. Lund, Sweden, Håkan Olssons Boktryckeri.

Mitchell, B. R., editor (1998). *International Historical Statistics: Europe, 1750-1993*. London: Macmillan Reference; New York: Stockton Press.

Mor, G. and Cardenas, I. (2010). The immune system in pregnancy: A unique complexity. *American Journal of Reproductive Immunology*, 63(6):425–433.

Neumayer, E. (2004). Recessions lower (some) mortality rates: Evidence from Germany. *Social Science & Medicine*, 58(6):1037–1047.

Odin, M. (1934). Sjukdomar och sjukdomsfrekvens i övre norrland särskilt med hänsyn til födans sammansättning (diseases and disease incidence in upper norrland with particular focus on the dietary composition). *Socialhygienisk undersökning i Västerbottens och Norrbottens län. Lund, Sweden: Håkan Olssons Boktryckeri.*

Pearce, J. and Langley-Evans, S. (2013). The types of food introduced during complementary feeding and risk of childhood obesity: A systematic review. *International Journal of Obesity*, 37:477–485.

Pelkowski, J. M. and Berger, M. C. (2004). The impact of health on employment, wages, and hours worked over the life cycle. *The Quarterly Review of Economics and Finance*, 44:102–121.

Pettersson-Lidbom, P. (2014). Midwives and maternal mortality: Evidence from a midwifery policy experiment in Sweden in the 19th century. Technical report, mimeo.

Richter, A. and Robling, P. O. (2013). Multigenerational effects of the 1918-19 influenza pandemic in Sweden. *Swedish Institute for Social Research*, (5).

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

Ruhm, C. J. (2000). Are recessions good for your health? *The Quarterly Journal of Economics*, 115(2):617–650.

Ruhm, C. J. (2003). Good times make you sick. *Journal of Health Economics*, 22(4):637–658.

Ruhm, C. J. (2004). Macroeconomic conditions, health and mortality. Technical report, National Bureau of Economic Research.

Schmidt, I. M., Jørgensen, M., and Michaelsen, K. F. (1995). Height of conscripts in Europe: Is postneonatal mortality a predictor? *Annals of Human Biology*, 22(1):57–67.

Scholte, R., van den Berg, G. J., and Lindeboom, M. (2012). Long-run effects of gestation during the Dutch hunger winter famine on labor market and hospitalization outcomes.

Skatteverket (2015). Sveriges församlingar genom tiderna:församlingar med förlossningsanstalter fram till 1947, http://www.skatteverket.se/privat/ folkbokforing/omfolkbokforing.html.

Sköld, P. (1996). From inoculation to vaccination: Smallpox in Sweden in the eighteenth and nineteenth centuries. *Population Studies*, 50(2):247–262.

Socialstyrelsen, S. (1938). *Levnadsvillkor och hushållsvanor i städer och industriorter omkring år 1933*. Sveriges officiella statistik.

Statistics Sweden (1935). Årsbok för Sveriges kommuner: Statistical yearbook of administrative districts of Sweden, 1931–1935.

Statistics Sweden (1997). Dödsorsaker 1995.

Statistics Sweden (2005). Multi-generation register 2005 – A description of contents and quality. *Örebro: Statistics Sweden*, pages 1–88.

The National Board of Health & Welfare (2014). Dödsorsaksregistret.

Thompson, O. (2014). Genetic mechanisms in the intergenerational transmission of health. *Journal of Health Economics*, 35:132–146.

Torell, U. (2013). När sockret blev vardagsmat. *Tidsskrift for kulturforskning*, 12(3-4).

Trannoy, A., Tubeuf, S., Jusot, F., and Devaux, M. (2010). Inequality of opportunities in health in France: A first pass. *Health Economics*, 19(8):921–938.

Van den Berg, G. J., Doblhammer, G., and Christensen, K. (2009). Exogenous determinants of early-life conditions, and mortality later in life. *Social Science & Medicine*, 68(9):1591–1598.

Van den Berg, G. J., Doblhammer-Reiter, G., and Christensen, K. (2011). Being born under adverse economic conditions leads to a higher cardiovascular mortality rate later in life: Evidence based on individuals born at different stages of the business cycle. *Demography*, 48(2):507–530.

Van den Berg, G. J., Lindeboom, M., and Portrait, F. (2006). Economic conditions early in life and individual mortality. *The American Economic Review*, pages 290–302.

Van den Berg, G. J. and Pinger, P. R. (2014). Transgenerational effects of childhood conditions on third generation health and education outcomes. *SOEP Paper*, (709):343–356.

Weber, M., Grote, V., Closa-Monasterolo, R., Escribano, J., Langhendries, J., Dain, E., Giovannini, M., Verduci, E., Gruszfeld, D., Socha, P., and Koletzko, B. (2014). Lower protein content in infant formula reduces BMI and obesity risk at school age: Follow-up of a randomized trial. *American Journal of Clinical Nutrition*, 99(5):1041–1051.

Weinehall, L., Westman, G., Hellsten, G., Boman, K., Hallmans, G., Pearson, T. A., and Wall, S. (1999). Shifting the distribution of risk: Results of a community intervention in a Swedish programme for the prevention of cardiovascular disease. *Journal of Epidemiology and Community Health*, 53(4):243–250.

Wisselgren, M. (2005). Att föda barn – från privat till offentlig angelägenhet. Technical report, Umeå University.

6 Figures

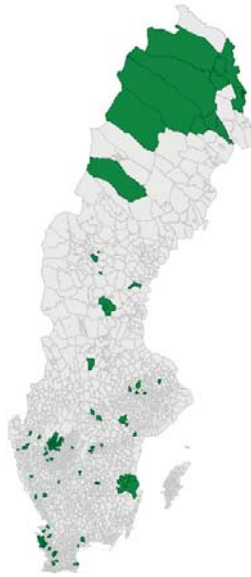


Figure 1: The location of the 133 parishes in the dataset.

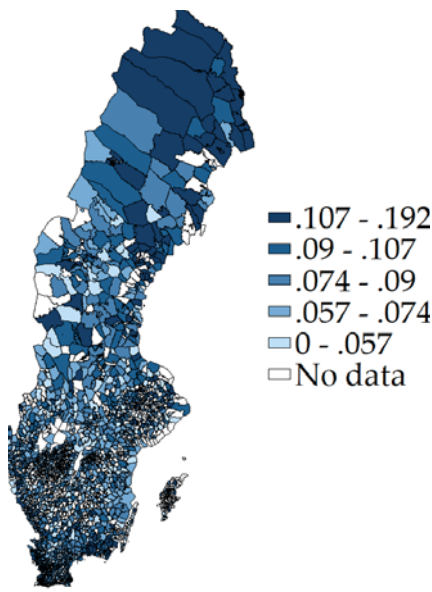


Figure 2: Average level of the infant mortality rate (IMR) between 1880 and 1919.

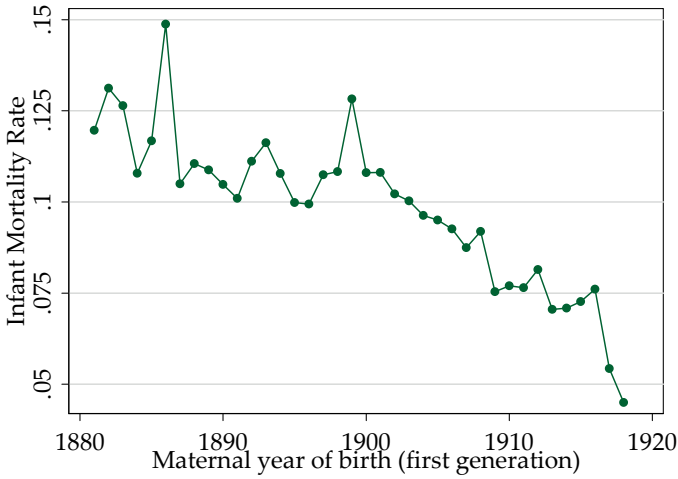
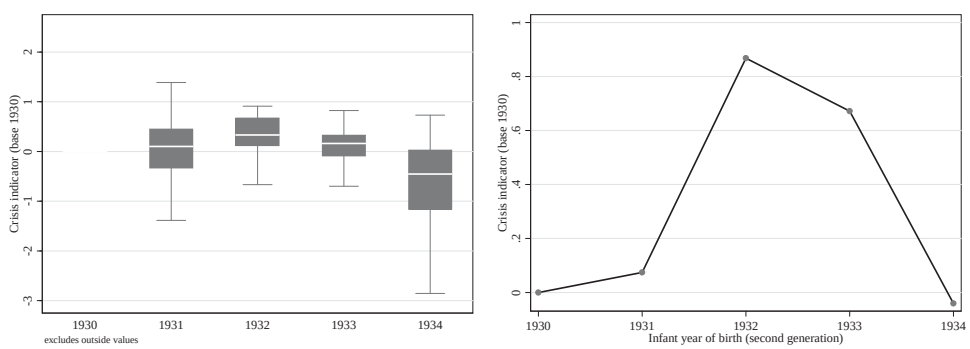


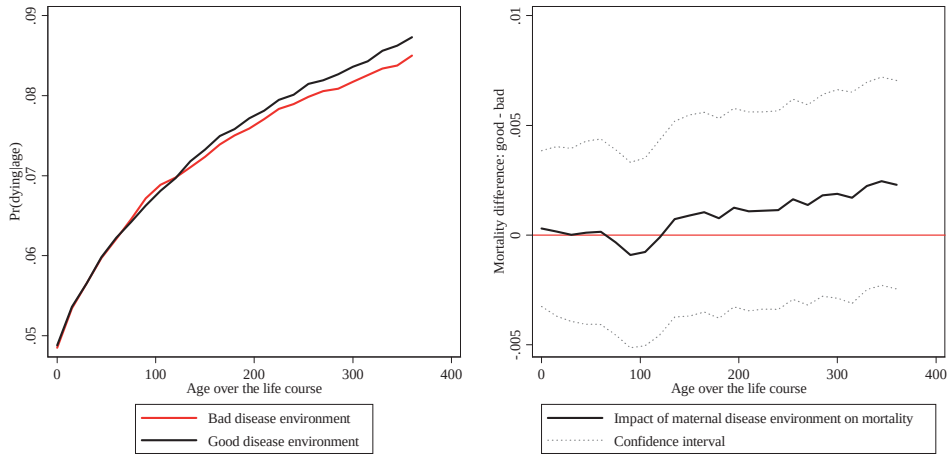
Figure 3: Average infant mortality rate (IMR) by maternal year of birth.



(a) Dynamics of some quartiles (25%, 50%, 75%)

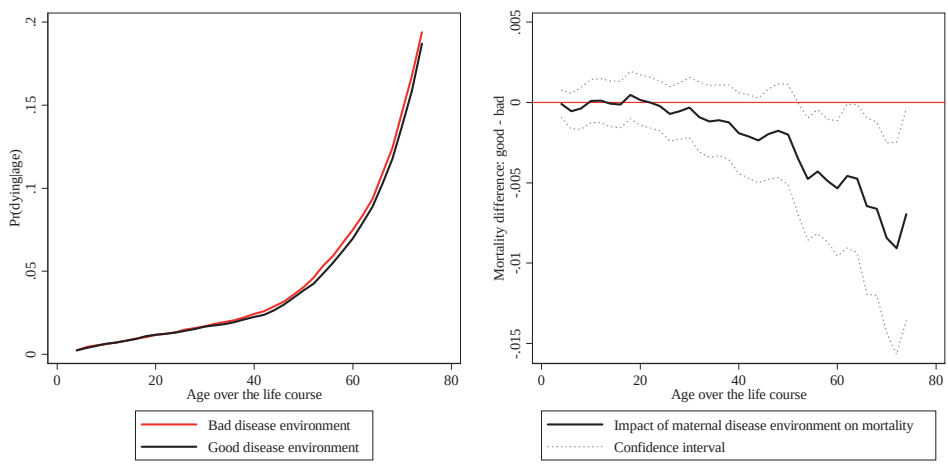
(b) Dynamics of the mean

Figure 4: Crisis indicator: average negative log deviation from 1930 tax revenues.



Note: The left panel depicts the cumulated hazard rate over the life course (bin: 15 days). The right panel depicts the difference of those cumulated hazard rates between good and bad maternal disease environment at birth.
Interval computed at the 90% confidence level. Standard errors computed based on 100 bootstrap replications.

Figure 5: Second generation mortality hazards in the first year of life.



Note: The left panel depicts the cumulated hazard rate over the life course (bin: two years). The right panel depicts the difference of those cumulated hazard rates between good and bad maternal disease environment at birth.
Interval computed at the 90% confidence level. Standard errors computed based on 100 bootstrap replications.

Figure 6: Second generation mortality hazards over the entire life cycle.

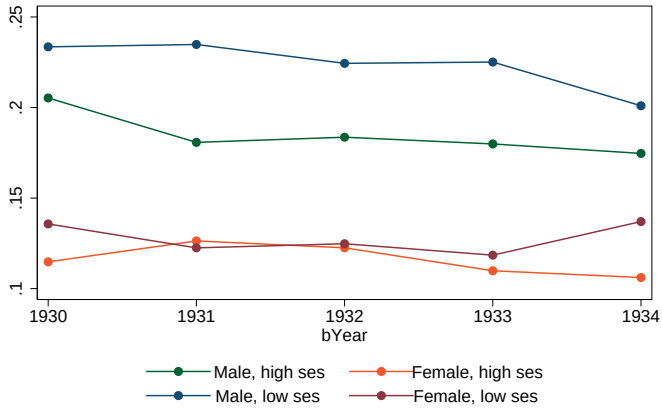


Figure 7: Cohort trends in mortality between 50 and 70, by sex and socioeconomic status (SES).

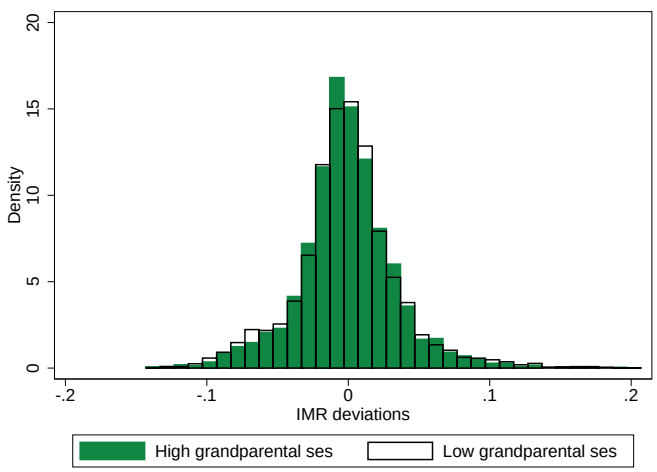


Figure 8: Distribution of the maternal disease environment by grandparental socioeconomic status (SES).

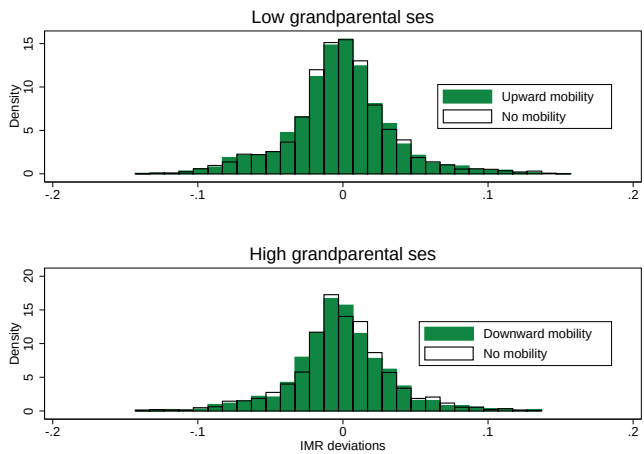


Figure 9: Distribution of the maternal disease environment by parental and grandparental socioeconomic status (SES).

7 Tables

Table 1: Descriptive statistics.

Variable	Mean	SD	Min	Max	Obs
Individual birth/death data					
IMR^j	0.10	0.05	0	1	25,010
$crisis^i$	0.05	0.20	-0.59	0.61	25,010
Female	0.48	0.50	0	1	25,010
Twin	0.03	0.16	0	1	25,010
Wedlock	0.89	0.31	0	1	25,005
Age mother (yrs)	29.13	6.66	13	50	25,007
Urban	0.21	0.41	0	1	25,010
Mortality 0-1 (baseline)	0.08	0.28	0	1	25,010
Mortality 1-50 (baseline)	0.08	0.27	0	1	22,940
Mortality 50-70 (baseline)	0.16	0.37	0	1	21,081
Mortality 70+ (baseline)	0.30	0.46	0	1	17,632
Census 1970					
Log labor income in 1970	8.41	3.56	0	13.37	18,566
Years of education	9.62	2.46	7.69	19	18,372
Death causes 1960-2013					
All-causes	0.41	0.49	0	1	22,739
Cancer (excl. lung/oral cavity)	0.08	0.28	0	1	22,739
Lung/oral cavity cancer	0.02	0.14	0	1	22,739
Respiratory diseases	0.04	0.19	0	1	22,739
CVD (no curable risk factors)	0.09	0.28	0	1	22,739
CVD (with curable risk factors)	0.03	0.17	0	1	22,739
Other circulatory system diseases	0.03	0.18	0	1	22,739
External causes/infections	0.03	0.18	0	1	22,739
Digestive/endocrine system (incl. diabetes)	0.01	0.12	0	1	22,739
Other symptoms	0.04	0.20	0	1	22,739
Not classified elsewhere	0.03	0.17	0	1	22,739

Table 2: Occupation of the household head according to the HISCO classification.

HISCO cat.	Description	Freq.	Percent	Cum.
0	Professional, technical and related workers	330	1.32	1.32
1		806	3.22	4.54
2	Administrative and managerial workers	517	2.07	6.61
3	Clerical and related workers	368	1.47	8.08
4	Sales workers	637	2.55	10.63
5	Service workers	439	1.76	12.38
6	Agricultural, animal husbandry and forestry workers, fishermen, hunters	8,913	35.64	48.02
7	Production and related workers, transport	2,291	9.16	57.18
8		1,324	5.29	62.47
9	Equipment operators and labourers	6,537	26.14	88.61
10	Unknown	2,848	11.39	100
	Total	25,010	100	

Table 3: Regression results: second generation mortality.

Dependent Variable: Mortality between Ages				
	0-1	1-50	50-70	70-
IMR^j	0.043	-0.012	0.141**	-0.045
	(0.046)	(0.037)	(0.062)	(0.077)
$crisis^i$	-0.009	0.015	-0.036*	-0.067**
	(0.015)	(0.013)	(0.019)	(0.028)
Female	-0.023***	-0.038***	-0.082***	-0.110***
	(0.004)	(0.003)	(0.005)	(0.007)
j's county of birth FE	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓
i's parish of birth FE	✓	✓	✓	✓
Baseline (%)	8.3	7.7	16.4	30.3
R-squared	0.017	0.019	0.023	0.029
N	25,010	22,940	21,081	17,632

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-county of birth level. Dep var: mortality in different phases of life (0-1; 1-50; 50-70; 70- y.o.) for the second generation.

Table 4: Regression results: second generation mortality (50-70). Robustness and effect heterogeneity.

		Dependent Variable: Mortality 50-70					
		(1)	(2)	(3)	(4)	(5)	(6)
A. All	IMR^j	0.141**	0.136**	0.136**	0.136**	0.135**	0.133**
		(0.062)	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
	R-squared	0.023	0.026	0.027	0.028	0.028	0.029
	N	21,081	21,081	21,081	21,081	21,081	21,081
B. Females	IMR^j	0.022	0.016	0.018	0.018	0.017	0.012
		(0.074)	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)
	R-squared	0.019	0.027	0.028	0.028	0.029	0.032
	N	10,561	10,561	10,561	10,561	10,561	10,561
C. Males	IMR^j	0.254***	0.246**	0.239**	0.243**	0.242**	0.241**
		(0.098)	(0.097)	(0.098)	(0.098)	(0.098)	(0.099)
	R-squared	0.020	0.026	0.028	0.028	0.028	0.032
	N	10,520	10,520	10,520	10,520	10,520	10,520
D. Mid-high SES	IMR^j	0.217**	0.214**	0.214**	0.215**	0.214**	0.211**
		(0.097)	(0.096)	(0.096)	(0.096)	(0.096)	(0.097)
	R-squared	0.028	0.035	0.035	0.035	0.035	0.038
	N	10,223	10,223	10,223	10,223	10,223	10,223
E. Low SES	IMR^j	0.079	0.079	0.079	0.077	0.076	0.076
		(0.077)	(0.077)	(0.077)	(0.078)	(0.078)	(0.078)
	R-squared	0.035	0.041	0.041	0.042	0.042	0.046
	N	10,858	10,858	10,858	10,858	10,858	10,858
F. Mid-high SES, females	IMR^j	0.064	0.057	0.057	0.062	0.060	0.048
		(0.115)	(0.115)	(0.115)	(0.115)	(0.115)	(0.115)
	R-squared	0.039	0.052	0.052	0.053	0.053	0.058
	N	5,111	5,111	5,111	5,111	5,111	5,111
G. Low SES, females	IMR^j	0.009	0.011	0.011	0.010	0.008	0.013
		(0.097)	(0.098)	(0.098)	(0.098)	(0.098)	(0.099)
	R-squared	0.035	0.048	0.048	0.049	0.049	0.056
	N	5,450	5,450	5,450	5,450	5,450	5,450
H. Mid-high SES, males	IMR^j	0.384***	0.379***	0.379***	0.383***	0.386***	0.382***
		(0.145)	(0.144)	(0.144)	(0.144)	(0.144)	(0.146)
	R-squared	0.034	0.045	0.045	0.046	0.046	0.053
	N	5,112	5,112	5,112	5,112	5,112	5,112
I. Low SES, males	IMR^j	0.159	0.166	0.166	0.166	0.163	0.155
		(0.124)	(0.125)	(0.125)	(0.126)	(0.126)	(0.127)
	R-squared	0.039	0.048	0.048	0.050	0.051	0.056
	N	5,408	5,408	5,408	5,408	5,408	5,408
i's ind. controls		✓	✓	✓	✓	✓	✓
j's county of birth FE		✓	✓	✓	✓	✓	✓
j's year of birth FE		✓	✓	✓	✓	✓	✓
i's parish of birth FE		✓	✓	✓	✓	✓	✓
j's county x time trends			✓	✓	✓	✓	✓
i's parent occupation FE				✓	✓	✓	✓
i's order of birth & twin dummies					✓	✓	✓
i's year/quarter of birth						✓	✓
i's mother age FE							✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Individual controls include: crisis indicator and female (specifications not split by sex only). Fixed effects on i's parent occupation included in specifications not split by SES only. Dep. var: mortality between 50-70 y.o. for the second generation; sample conditional on survival until 50 y.o.

Table 5: Regression results: mortality by cause of death. Robustness and effect heterogeneity.

		All	By Sex		By SES		By Sex x Socio-Economic Status			
			Females	Males	Mid-High	Low	Females, Mid-High SES	Females, Low SES	Males, Mid-High SES	Males, Low SES
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. All-cause mortality	IMR^j	0.133**	0.012	0.241**	0.211**	0.076	0.048	0.013	0.382***	0.155
		(0.062)	(0.074)	(0.099)	(0.097)	(0.078)	(0.115)	(0.099)	(0.146)	(0.127)
	Prevalence	0.164	0.122	0.205	0.150	0.176	0.116	0.128	0.185	0.224
B. Cardiovascular diseases (no preventable/curable risk factors)	IMR^j	0.066*	0.039	0.090	0.126*	0.028	0.054	0.023	0.202*	0.038
		(0.035)	(0.032)	(0.063)	(0.066)	(0.035)	(0.054)	(0.041)	(0.119)	(0.056)
	Prevalence	0.039	0.021	0.056	0.036	0.042	0.020	0.022	0.051	0.061
C. Cardiovascular diseases (preventable/curable risk factors)	IMR^j	-0.002	-0.024	0.016	-0.006	0.001	-0.056**	0.002	0.047	-0.001
		(0.017)	(0.023)	(0.025)	(0.021)	(0.025)	(0.025)	(0.037)	(0.035)	(0.037)
	Prevalence	0.011	0.005	0.016	0.009	0.013	0.004	0.007	0.013	0.019
D. Other causes:										
Diseases of the digestive/ endocrine system incl. diabetes	IMR^j	0.016	0.016	0.018	0.036**	-0.001	0.039	0.001	0.046*	-0.003
		(0.011)	(0.016)	(0.017)	(0.018)	(0.014)	(0.026)	(0.016)	(0.025)	(0.024)
	Prevalence	0.007	0.005	0.008	0.006	0.007	0.006	0.005	0.006	0.010
Lung & oral cavity cancer	IMR^j	-0.027*	-0.031*	-0.023	-0.026	-0.029	-0.024	-0.028	-0.022	-0.029
		(0.015)	(0.019)	(0.022)	(0.021)	(0.021)	(0.027)	(0.024)	(0.029)	(0.033)
	Prevalence	0.009	0.007	0.012	0.009	0.010	0.006	0.007	0.011	0.013
Cancer (excl. lung & oral cavity cancer)	IMR^j	0.016	-0.038	0.073*	0.042	-0.015	-0.014	-0.051	0.113	0.029
		(0.026)	(0.037)	(0.039)	(0.046)	(0.034)	(0.057)	(0.053)	(0.070)	(0.045)
	Prevalence	0.039	0.041	0.037	0.038	0.040	0.041	0.041	0.036	0.039
Symptoms/signs not elsewhere classified	IMR^j	0.027	0.028	0.027	-0.042**	0.082**	-0.041	0.090*	-0.048	0.085*
		(0.022)	(0.030)	(0.032)	(0.020)	(0.036)	(0.027)	(0.053)	(0.032)	(0.049)
	Prevalence	0.010	0.009	0.011	0.010	0.011	0.008	0.010	0.011	0.012
External causes/ infections/par. diseases	IMR^j	-0.000	0.019	-0.021	0.019	-0.016	0.058*	-0.018	-0.025	-0.011
		(0.021)	(0.018)	(0.037)	(0.025)	(0.032)	(0.034)	(0.017)	(0.039)	(0.057)
	Prevalence	0.013	0.008	0.019	0.012	0.014	0.008	0.008	0.016	0.021
Other circulatory system diseases (excl. B. and C.)	IMR^j	0.037	0.001	0.072	0.032	0.041	-0.015	0.016	0.086*	0.071
		(0.025)	(0.016)	(0.047)	(0.024)	(0.040)	(0.020)	(0.026)	(0.047)	(0.074)
	Prevalence	0.012	0.008	0.017	0.011	0.013	0.007	0.008	0.016	0.019
Respiratory diseases	IMR^j	-0.013	-0.024	-0.012	0.007	-0.027	0.002	-0.035	-0.001	-0.034
		(0.014)	(0.019)	(0.021)	(0.023)	(0.018)	(0.033)	(0.023)	(0.031)	(0.029)
	Prevalence	0.011	0.009	0.013	0.010	0.011	0.008	0.009	0.012	0.014
All other causes	IMR^j	0.015	0.025	0.002	0.021	0.016	0.041	0.015	-0.012	0.015
		(0.018)	(0.026)	(0.025)	(0.032)	(0.020)	(0.053)	(0.023)	(0.038)	(0.034)
	Prevalence	0.013	0.010	0.016	0.011	0.015	0.008	0.011	0.013	0.019
i's ind. controls		✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county of birth FE		✓	✓	✓	✓	✓	✓	✓	✓	✓
j's year of birth FE		✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county x time trends		✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parent occupation FE		✓	✓	✓						
i's order of birth & twin dummies		✓	✓	✓	✓	✓	✓	✓	✓	✓
i's year/quarter of birth FE		✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parish of birth		✓	✓	✓	✓	✓	✓	✓	✓	✓
i's mother age FE		✓	✓	✓	✓	✓	✓	✓	✓	✓
N		21,081	10,561	10,520	10,223	10,858	5,111	5,450	5,112	5,408

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-county of birth level. Dep. var: death causes for the second generation (conditional on surviving until 50 y.o.). Category "All other causes" is a residual category for all death causes not included in the other reported categories. Overall effect in col. 1, split by sex (2-3), split by SES (4-5), interaction sex*SES (6-9).

Table 6: Regression results: earnings in 1970.

		Dependent Variable: Log Labor Income in 1970					
		(1)	(2)	(3)	(4)	(5)	(6)
A. All	IMR^j	-0.986*	-0.914*	-0.907*	-0.921*	-0.915*	-0.920*
		(0.539)	(0.546)	(0.544)	(0.542)	(0.542)	(0.541)
	R-squared	0.208	0.211	0.213	0.213	0.214	0.215
	N	18,566	18,566	18,566	18,566	18,566	18,566
B. Females	IMR^j	-2.183**	-2.147**	-2.121**	-2.108**	-2.083**	-2.073**
		(1.036)	(1.052)	(1.050)	(1.050)	(1.049)	(1.048)
	R-squared	0.027	0.036	0.039	0.039	0.040	0.043
	N	9,104	9,104	9,104	9,104	9,104	9,104
C. Males	IMR^j	-0.002	0.095	0.108	0.081	0.088	0.082
		(0.425)	(0.420)	(0.422)	(0.423)	(0.421)	(0.422)
	R-squared	0.024	0.031	0.033	0.034	0.035	0.040
	N	9,462	9,462	9,462	9,462	9,462	9,462
D. Mid-high SES	IMR^j	-0.354	-0.409	-0.409	-0.424	-0.401	-0.481
		(0.759)	(0.763)	(0.763)	(0.760)	(0.758)	(0.757)
	R-squared	0.227	0.232	0.232	0.233	0.233	0.236
	N	10,034	10,034	10,034	10,034	10,034	10,034
E. Low SES	IMR^j	-1.480*	-1.310	-1.310	-1.359*	-1.377*	-1.333*
		(0.802)	(0.808)	(0.808)	(0.806)	(0.805)	(0.803)
	R-squared	0.204	0.209	0.209	0.210	0.211	0.214
	N	8,532	8,532	8,532	8,532	8,532	8,532
F. Mid-high SES, females	IMR^j	-1.667	-1.692	-1.692	-1.656	-1.563	-1.632
		(1.477)	(1.479)	(1.479)	(1.474)	(1.476)	(1.474)
	R-squared	0.048	0.062	0.062	0.063	0.065	0.070
	N	4,926	4,926	4,926	4,926	4,926	4,926
G. Low SES, females	IMR^j	-2.426	-2.236	-2.236	-2.258	-2.281	-2.103
		(1.492)	(1.525)	(1.525)	(1.523)	(1.522)	(1.516)
	R-squared	0.046	0.061	0.061	0.062	0.063	0.070
	N	4,178	4,178	4,178	4,178	4,178	4,178
H. Mid-high SES, males	IMR^j	0.326	0.325	0.325	0.298	0.299	0.218
		(0.487)	(0.488)	(0.488)	(0.489)	(0.490)	(0.493)
	R-squared	0.044	0.056	0.056	0.057	0.058	0.063
	N	5,108	5,108	5,108	5,108	5,108	5,108
I. Low SES, males	IMR^j	-0.390	-0.159	-0.159	-0.185	-0.186	-0.164
		(0.690)	(0.675)	(0.675)	(0.677)	(0.676)	(0.677)
	R-squared	0.044	0.059	0.059	0.060	0.061	0.073
	N	4,354	4,354	4,354	4,354	4,354	4,354
i's ind. controls		✓	✓	✓	✓	✓	✓
j's county of birth FE		✓	✓	✓	✓	✓	✓
j's year of birth FE		✓	✓	✓	✓	✓	✓
i's parish of birth FE		✓	✓	✓	✓	✓	✓
j's county x time trends			✓	✓	✓	✓	✓
i's parent occupation FE				✓	✓	✓	✓
i's order of birth & twin dummies					✓	✓	✓
i's year/quarter of birth						✓	✓
i's mother age FE							✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Individual controls include: crisis indicator and female (specifications not split by sex only). Fixed effects on i's parent occupation included in specifications not split by SES only. Dep. var: log labor income measured in 1970 (from Census 1970).

Table 7: Regression results: employment in 1970.

	All			Females			Males		
	Any (1)	Fulltime (2)	Part-time (3)	Any (4)	Fulltime (5)	Part-time (6)	Any (7)	Fulltime (8)	Part-time (9)
IMR^j	-0.110*	-0.008	-0.102**	-0.355***	-0.142	-0.213**	0.080	0.086	-0.006
	(0.061)	(0.061)	(0.047)	(0.111)	(0.109)	(0.093)	(0.054)	(0.059)	(0.028)
$crisis^i$	-0.015	-0.006	-0.009	0.018	0.038	-0.020	-0.033	-0.035	0.003
	(0.025)	(0.025)	(0.020)	(0.044)	(0.042)	(0.039)	(0.022)	(0.025)	(0.012)
Female	-0.332***	-0.571***	0.240***						
	(0.006)	(0.006)	(0.005)						
i's ind. controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county x time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parent occupation FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's order of birth & twin dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's year/quarter of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parish of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's mother age FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline (%)	76.7	63.3	13.4	59.5	33.8	25.7	93.3	91.6	1.6
R-squared	0.185	0.381	0.141	0.084	0.110	0.040	0.069	0.064	0.023
N	18,566	18,566	18,566	9,104	9,104	9,104	9,462	9,462	9,462

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Dep. var: probability of being employed in any job, in a full time job, in a part time job, respectively. Split by sex: columns 4-6: females; columns 7-9: males.

Table 8: Regression results: years of schooling.

	All	By Sex		By Socio-Economic Status		By Sex x Socio-Economic Status			
		Females	Males	Mid-High	Low	Females, Mid-High SES	Females, Low SES	Males, Mid-High SES	Males, Low SES
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMR^j	0.258	0.460	0.039	0.545	0.074	1.044	-0.023	-0.002	0.281
	(0.382)	(0.511)	(0.542)	(0.666)	(0.487)	(0.950)	(0.634)	(0.875)	(0.712)
crisis^i	-0.000	0.219	-0.179	0.007	-0.039	0.193	0.056	-0.048	-0.137
	(0.137)	(0.200)	(0.208)	(0.214)	(0.186)	(0.325)	(0.258)	(0.307)	(0.286)
Female	-0.145***			-0.068	-0.200***				
	(0.034)			(0.051)	(0.046)				
i's ind. controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county x time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parent occupation FE	✓	✓	✓						
i's order of birth & twin dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's year/quarter of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parish of birth	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's mother age FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline	9.588	9.518	9.657	9.698	9.485	9.664	9.382	9.731	9.587
R-squared	0.100	0.117	0.115	0.116	0.060	0.138	0.094	0.148	0.081
N	20,518	10,139	10,379	9,919	10,599	4,903	5,236	5,016	5,363

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Dep. var: years of education in 1970 for the second generation.

Table 9: Regression results: employment in public services in 1970.

	All	By Sex		By Socio-Economic Status		By Sex x Socio-Economic Status			
		Females	Males	Mid-High	Low	Females, Mid-High SES	Females, Low SES	Males, Mid-High SES	Males, Low SES
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMR^j	-0.076	-0.158*	-0.009	-0.026	-0.105	-0.046	-0.252**	0.027	0.015
	(0.050)	(0.083)	(0.063)	(0.080)	(0.067)	(0.132)	(0.110)	(0.094)	(0.088)
crisis^i	0.000	0.019	-0.013	-0.007	-0.006	0.004	0.005	-0.013	-0.008
	(0.020)	(0.034)	(0.025)	(0.030)	(0.030)	(0.052)	(0.048)	(0.034)	(0.040)
Female	0.073***			0.086***	0.062***				
	(0.005)			(0.007)	(0.007)				
i's ind. controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
j's county x time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parent occupation FE	✓	✓	✓						
i's order of birth & twin dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's year/quarter of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's parish of birth	✓	✓	✓	✓	✓	✓	✓	✓	✓
i's mother age FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline (%)	12.3	16	8.6	13.2	11.2	17.4	14.5	9.1	8.1
R-squared	0.047	0.050	0.052	0.056	0.050	0.064	0.080	0.075	0.062
N	18,566	9,104	9,462	10,034	8,532	4,926	4,178	5,108	4,354

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Dep. Var: being employed in public services in 1970, overall effect (1), split by sex (2-3), split by SES (4-5) and interactions sex*SES (6-9).

Table 10: Assessing the potential selection effects: selection into fertility and in utero.

Dependent var:	crisis^i	Female	Twin	Wedlock	Age mother (yrs)	SES (low=1)
	(1)	(2)	(3)	(4)	(5)	(6)
IMR^j	-0.011	0.034	0.010	-0.063	-0.248	0.076
	(0.016)	(0.071)	(0.028)	(0.051)	(0.242)	(0.079)
j's county of birth FE	✓	✓	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓	✓	✓
j's county x year of birth FE	✓	✓	✓	✓	✓	✓
i's parent occupation FE	✓	✓	✓	✓	✓	
i's order of birth FE	✓	✓	✓	✓	✓	✓
i's year/quarter of birth FE	✓	✓	✓	✓	✓	✓
i's parish of birth	✓	✓	✓	✓	✓	✓
i's mother age FE	✓	✓	✓	✓		✓
Baseline	0.05	0.48	0.03	0.89	28.8	0.5
R-squared	0.650	0.012	0.051	0.287	0.970	0.167
N	25,009	25,009	25,009	25,004	25,009	25,009

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-county of birth level. Dep. var: various characteristics at birth of the second generation. The regression results are indicative of whether the IMR variable relates to fertility decisions.

Table 11: Regression of the infant mortality rate (IMR) on potential drivers (on regional level).

Dep. var: Ln(IMR)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Logarithm of diphtheria morbidity rate	0.0186***								0.0158**
	(0.00676)								(0.0068)
2. Logarithm of scarlet fever morbidity rate		0.0140							0.0143*
		(0.00851)							(0.0084)
3. Logarithm of respiratory disease morbidity rate			0.00926**						0.00702*
			(0.00405)						(0.00406)
4. Logarithm of share of vaccinations per capita				-0.00489					-0.00491
				(0.00321)					(0.00318)
5. Logarithm of farmhand wage					0.00298				0.00193
					(0.00588)				(0.00588)
6. Log number of pharmacies						200793.5*			159026.9
						(102522.3)			(103237.3)
7. Logarithm of midwives per females							-0.00225		-0.000573
							(0.00849)		(0.00842)
8. Log number of doctors								-0.0181*	-0.00537
								(0.00993)	(0.00367)
F-test of joint influence of disease environment (coefficients 1-3), *p* value in parentheses								3.99	(0.008)
F-test of joint influence of health care accessibility and living standards (coefficients 4-8), *p* value in parentheses								1.50	(0.188)
Region FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Region specific trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	550	550	550	550	550	550	550	550	550

* p<0.1, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Note: the infant mortality rate (IMR) is measured on the regional level. Data ranges from 1889 to 1910 (22 years, 86% of the first generation mothers are born in this period) and comprises all 25 regions of Sweden.

Table 12: Assessing the the underlying channel behind the second generation mortality: infant mortality rate (IMR) in utero versus in early life.

Dependent Variable: Mortality between Ages				
	0-1	1-50	50-70	70-
IMR^j	0.043	-0.012	0.141**	-0.045
	(0.046)	(0.037)	(0.062)	(0.077)
Baseline (%)	8.3	7.7	16.4	30.3
R-squared	0.017	0.019	0.023	0.029
N	25,010	22,940	21,081	17,632
IMR^j (in-utero sample)	-0.034	-0.016	0.158*	-0.099
	(0.058)	(0.063)	(0.092)	(0.117)
Baseline (%)	7.3	7.3	16.3	30
R-squared	0.022	0.024	0.035	0.036
N	11,035	10,233	9,446	7,905
IMR in-utero	-0.016	0.009	0.151*	0.128
	(0.047)	(0.051)	(0.079)	(0.104)
Baseline (%)	7.3	7.3	16.3	30
R-squared	0.022	0.024	0.035	0.036
N	11,035	10,233	9,446	7,905
j's county of birth FE	✓	✓	✓	✓
j's year of birth FE	✓	✓	✓	✓
i's parish of birth FE	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses clustered at the mother j year of birth-county of birth level. Robustness check: effect of IMR on mortality in 4 different phases of life (0-1, 1-10, 50-70, 70- y.o. in cols 1-4). First panel: benchmark (IMR measured in the year of birth of the mother); second panel: IMR in the year of birth of the mother on the subsample for which we have also IMR in-utero information; third panel: IMR in utero.

Table 13: Regression results, first generation mortality between 50 and 70.

	All	**By SES**	
		Mid-High	Low
	(1)	(2)	(3)
IMR^j	-0.035	0.012	-0.083
	(0.066)	(0.096)	(0.085)
mother_nameSES	-0.008		
	(0.008)		
j's county of birth FE	✓	✓	✓
j's year of birth FE	✓	✓	✓
i's parish of birth	✓	✓	✓
R-squared	0.017	0.030	0.026
N	15,733	7,105	8,628

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses clustered at the mother j year of birth-region of birth level. Dep. var: mortality between 50-70 y.o. for the first generation; sample conditional on survival until 50 y.o.

Table 14: Cox proportional hazard regressions for mothers.

		Dependent Variable: 1st Generation Mortality			
		(1)	(2)	(3)	(4)
A. All	IMR^j	0.176	0.166	0.154	0.143
		(0.186)	(0.185)	(0.184)	(0.184)
	N.	16,961	16,961	16,961	16,961
B. Mid-high SES	IMR^j	-0.136	-0.154	-0.154	-0.172
		(0.258)	(0.259)	(0.259)	(0.260)
	N.	7,580	7,580	7,580	7,580
C. Low SES	IMR^j	0.465**	0.479**	0.479**	0.460**
		(0.222)	(0.221)	(0.221)	(0.220)
	N.	9,381	9,381	9,381	9,381
j's individual characteristics		✓	✓	✓	✓
j's county of birth FE		✓	✓	✓	✓
j's year of birth FE		✓	✓	✓	✓
Children's parish of birth		✓	✓	✓	✓
j's county x year of birth FE			✓	✓	✓
Household head occupation FE				✓	✓
N. of children FE					✓

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. Household head occupation FE included just in the specifications not split by SES (panel A). Dep. var: mortality in the first generation (life duration in days). Cox regression (coefficients reported).