The Economy, the Ghost in Your Gene and the Escape from Premature Mortality

by

Dora L. Costa, UCLA and NBER
Lars Olov Bygren, Karolinksa Institutet and Umeå Universitet
Benedikt Graf, Sustento Group
Martin Karlsson, CINCH, University of Duisberg Essen and IZA
Joseph Price, BYU and NBER

June 13, 2025

We have benefitted from the comments of Matthew Kahn and of participants at the 2023/24 Mackintosh Lecture at Queen's University, the 2023 meetings of the Western Economics Association, and the 2022 Essen Health Conference. We thank the many undergraduates in UCLA's HIGGS lab who contributed to this project. Dora Costa gratefully acknowledges the use of facilities and resources at the California Center for Population Research, UCLA, which is supported in part by NICDH grant P2C HD041022. Earlier findings from this project were presented at Stanford and Vanderbilt.

The Economy, The Ghost in Your Gene and the Escape from Premature Mortality

Abstract

Explanations for the West's escape from premature mortality have focused on chronic malnutrition or income and on public health or state capacity. We argue that by ignoring the multigenerational effects of variance in ancestors' harvests, we are underestimating the contribution of modern economic growth to the escape from early death at older ages. Using two multigenerational datasets for Sweden, we show that grandsons' longevity was strongly linked to spatial shocks in paternal grandfathers' yearly harvest variability when agricultural productivity was low and market integration was limited. We reason that an epigenetic mechanism is the most plausible explanation for our findings. We posit that the removal of trade barriers, improvements in transportation, and agricultural innovation reduced harvest variability. We contend that for older Swedish men (but not women) who reached age 40 between 1870 and 1919 this reduction was as important as decreasing contemporaneous infectious disease rates and more important than eliminating exposure to poor harvests in-utero. Decreasing contemporaneous infectious disease rates dominated reductions in paternal grandfathers' harvest variability for men reaching age 40 between 1920 and 1949.

JEL Classifications: I15, J11, N33

Corresponding author
Dora L. Costa
UCLA Department of Economics
9272 Bunche Hall
Los Angeles, CA 90095-1477
and NBER
costa@econ.ucla.edu

The past was a dangerous place. In Europe before the 1780s, the majority of households derived their income from agriculture and local harvest failures led to outbreaks of famine because of low agricultural productivity, limited storage, high transportation costs and trade barriers within countries (Dribe et al., 2015; Persson, 1999; Floud et al., 2011). Even after the elimination of crisis mortality, year to year variability in the food supply remained high throughout the nineteenth century (Persson, 1999; Campbell and Ó Gráda, 2011). Harvest variability remains a problem in developing countries[1] where the poor are particularly vulnerable,[2] and has become a concern in developed countries because of climate change. High economic uncertainty in turn has been linked to low economic activity (see Fernández-Villaverde and Guerrón-Quintana (2020) for a review).

We investigate the implications of yearly harvest swings on the older age longevity of later generations using two multigenerational databases (one newly constructed) for Sweden, the only country with detailed and reliable historical local harvest series and vital records. We study birth cohorts whose grandparents reached adulthood when the scope of adaptation to bad weather and poor harvests was limited. These cohorts experienced the nineteenth and twentieth century increases in longevity at older ages, declining infectious disease rates and the cardiovascular disease epidemic of the first half of the twentieth century.[3]

Our analyses shed new light on Europe's escape from premature death at older ages.

[1]The coefficient of variation on maize yields is high in parts of Brazil, Africa, India and Mexico and variation in rice yields is high in Brazil and central India (Ray et al., 2015).

[2]Exposure to both aggregate and idiosyncratic risk is greater among poorer household in developing countries (Samphantharak and Townsend, 2018),

[3]Conditional on surviving to age 40, among Swedes born in 1801, 88% of men and 83% of women were dead by age 81. A half century later the respective figures were 75 and 71%. Another half century later the respective figures were 68% and 49% (HMD. Human Mortality Database, 2024).

Prior researchers have pointed to increases in food levels or incomes (Floud et al., 2011) and reductions in infectious disease or state capacity (Preston, 1975; Easterlin, 1996; Deaton, 2013). We argue that a third factor, particularly for men, is ancestral variance in food levels. We are not discounting the role of chronic malnutrition (Floud et al., 2011), exhibited in shorter heights and lighter weights, in the escape from premature mortality, nor are we minimizing the contributions of scientific knowledge and public health investments in clean water (Alsan and Goldin, 2019; Cutler and Miller, 2005; Anderson et al., 2022), to the control of infectious disease. Rather, we argue that by ignoring the multigenerational effects of variance in ancestors' food levels, we are underestimating the contributions of modern economic growth to the escape from premature mortality.

We argue that a biological mechanism links longevity to ancestors' spatial shocks in yearly harvest variability when agricultural productivity was low and market integration was limited. We posit that the removal of trade barriers, improvements in transportation, and agricultural innovation reduced harvest variability. We show that the elimination of yearly harvest swings was more important than reductions in contemporaneous infectious disease rates for older Swedish men (but not women) who reached age 40 before 1920.

The long-run health implications of low nutritional *levels* are extensively documented. In-utero exposure to famines and malnutrition is associated with children's poor health at birth and in adulthood (Forsdahl, 1977; Barker, 1992, 1994; Lumey and van Poppel, 2013; Heijmans et al., 2008) and to grandchildren's health (Veenendaal et al., 2013). Less is known about health effects of *variability* in nutrition. Fluctuations in weight adjusted for height are associated with later life cardiovascular disease independent of weight for height levels, e.g. (Almuwaqqat et al., 2024). Two studies examine longer-run relationships. In

one study, later life cardiovascular and cerebrovascular disease are more strongly linked to in-utero variance in food availability than to low in-utero food levels (Bygren et al., 2000). The other study associates changes in paternal grandmothers' harvest conditions prior to puberty with granddaughters' cardiovascular mortality at older ages (Bygren et al., 2014).

We hypothesize that ancestral variance in food levels operates through an epigenetic mechanism, an inheritable modification of gene function, during the slow growing years prior to puberty.[4] We suspect that variance is a trigger for epigenetic processes because uncertainty in our evolutionary past required adaptability which is achieved by cells' entry into dormancy or quiescence, thus preserving cells for homeostasis and repair. Transitions in and out of dormancy require many regulatory layers, including gene regulation through micro-RNAs (Özgüldez and Bulut-Karslioğlu, 2024), Micro-RNAs have been implicated in epigenetic inheritance outside the germline (Yang et al., 2022; Reddy et al., 2021).

A distinguishing feature of our study from previous work providing proof of concept for transgenerational epigenetic inheritance in humans is our emphasis on harvest swings rather than levels. The three sets of transgenerational studies of human populations which examined grandchildren's outcomes and grandparents' food shocks during the vulnerable, slow growth period before puberty used data from the parish of Överkalix, Sweden (Kaati et al., 2007; Bygren et al., 2001), the county of Uppsala in Sweden (Vågerö et al., 2022, 2018), and Germany (Van den Berg and Pinger, 2016) to find that over-feeding of the paternal grandfather led to worse later-life outcomes among male-line grandsons, whereas under-feeding produces better outcomes. The Överkalix and Germany studies

[4]For a review of the biological processes, see Ryan and Kuzawa (2020). One of the first references in the economics literature to epigenetics was in Janet Currie's Richard T. Ely lecture (Currie, 2011).

found a similar association for paternal grandmothers' conditions and granddaughters' outcomes. Because of study designs, it is unclear whether high levels of food intake or variability in food intake leads to greater mortality among grandchildren.

The richness of our genealogical data enables us to implement empirical strategies which better control for place, cohort and family effects compared to prior research. Identification in the Överkalix and Germany studies was from grandparents' birth year. In the Uppsala studies, it was from the interaction between grandparents' birth year and county. Our identification also is from grandparents' birth year and county interactions but we control for harvest swings at all other ancestral ages between 0 and 20 and own and grandparents birth year and place fixed effects. We control for unobserved family effects by looking at second cousins, examining the descendants of grandfathers who were brothers where one experienced harvest swings during the epigenetically sensitive years and the other did not. In addition, we examine interactions with own and fathers' in-utero conditions, as proxied by semester of birth.

Another distinguishing feature of our paper is our ability to determine whether harvest variability left a lasting epigenetic legacy (the "ghost in your gene") on long-run population level mortality trends. Past studies focused on special populations (e.g. tiny, isolated Överkalix close to Arctic Circle and Germany during the WWI famine). We study earlier cohorts where over a *third* of grandparents experienced harvest variability during their slow growing years. By focusing on early cohorts we can parse out the contributions of declines in ancestral harvest variability and decreases in contemporaneous infectious disease rates to the escape from premature older age mortality.

The paper begins by discussing epigenetic and socioeconomic inheritance and by pre-

senting new evidence documenting a shift circa 1840 in the extent of harvest swings across Sweden and in the ability of the population to adapt to droughts and harvest swings. Our findings contribute to the literature analyzing trends in the scope of adaptation to climate change (Bleakley and Hong, 2017; Barreca et al., 2016; Kahn, 2010). We then describe our data and our methods. The results section shows that grandsons, but not granddaughters, descended from paternal grandfathers experiencing harvest swings during their slow growth years before puberty and reaching adulthood prior to 1840, were shorter lived than grandsons whose paternal grandfathers did not experience such swings. The mechanisms section examines alternative hypotheses, allowing us to rule out interpretations other than epigenetics, and investigates whether the level of economic development mediates epigenetic transmission. In the final section of the paper we demonstrate that in Sweden the ten year male birth cohorts most affected by ancestral harvest swings could have gained almost one and a half years of life at age 40 in the absence of paternal grandfathers' harvest swings. Eliminating harvest variability was as important for life expectancy as decreasing contemporaneous infectious rates for Swedish male (but not female) cohorts who reached age 40 between 1870 and 1919 and was more important than eliminating poor in-utero harvests. Paternal grandfathers' harvest swings were not as important as reductions in contemporaneous infectious disease rates for cohorts reaching age 40 between 1920 and 1949.

1 Background

There are two challenges in identifying epigenetic transmission using circumstantial evidence. The first is distinguishing epigenetic transmission from other forms of transmission such as socioeconomic, cultural or genetic. We will review the sex-specificity of epigenetic transmission and the importance of critical ancestral ages. We will show that sex-specific differences in the transmission of socioeconomic status and life-span were small for Swedish cohorts born in the late 1700s and in the 1800s.

The second challenge is obtaining data on harvest swings. We will demonstrate that harvest swings were more common prior to 1840 compared to after 1840 and that prior to 1840 the scope of adaptation to harvest swings was more limited.

1.1 Ancestral Inheritance

Males may be particularly sensitive to epigenetic transmission. The Y chromosome contains sequences crucial for spermatogenesis. The sperm epigenome is responsive to environmental influences through noncoding RNA (Chen et al., 2016; Nätt and Öst, 2020), which may in turn regulate DNA methylation and autosomal gene expression (Reddy et al., 2021).

Epigenetic markings transmitted through the Y chromosome imply an association between paternal grandfathers' harvest swings and grandsons' health at older agees because the Y chromosome is transmitted only through the male line. There should be no relationship between granddaughters' health and paternal grandfathers' harvest swings.

Epigenetic markings transmitted through the X chromosome suggest an association

between paternal grandmothers' harvest swings and granddaughters', but not grandsons', health. However, granddaughters may be less susceptible to epigenetic transmission if epigenetic marks are more likely to be erased when two X chromosomes pair and recombine during meiosis.[5] We cannot trace X chromosomes transmitted from maternal grandfathers and grandmothers.

Although epigenetic transmission can happen at any age (Murphy et al., 2016) and in a brief time frame Ryan and Kuzawa (2020), epidemiological evidence has identified the slow growth period just before puberty as sensitive to changes in the epigenome (Pembrey et al., 2006). This is the period when the first viable pools of spermatocytes and ovum emerge and when changes to the reproductive tract may be permanent.

A priori, it is unclear what, if any, type of change should be expected across generations under an epigenetic hypothesis. The inheritance of epigenetic marks across generations ranges from stability to total reset and encompasses even epigenetic recall, a facilitated response in descendants requiring an environmental inducer (Jablonka and Raz, 2009). Kaati et al. (2007) and Vågerö et al. (2018) documented generational skipping when the grandfather's stress occurred in the slow growth period before puberty.

The epigenetic hypothesis predicts that epigenetic inheritance is more likely when in-utero conditions are poor. Animal studies with maternal nutritional supplementation have reversed epigenetic signatures (Bernal et al. 2013; Dolinoy et al. 2007) and because interactions between imprinted genes on the X and Y chromosomes control placental size and thus fetal nourishment (Cunningham and Eghbali 2018). Epigenetic erasure and repro-

[5] Bygren et al. (2014) have suggested that because the X and Y chromosomes do not pair and recombine in men the epigenetic signal may remain intact in the father's generation to be inherited by the grandchildren via the gametes.

gramming occur in early and late gestation, respectively. Thus, the disruption of erasure or reprogramming because of poor nutrition could influence the interaction between epigenetic inheritance and the intrauterine environment (Baxter and Drake 2019).

The slow growth period prior to puberty should not be a critical age for non-epigenetic inheritance, nor should Sweden historically have had sex-specificity of socioeconomic, cultural or genetic transmission. Correlations with parents' life-spans were similar between sons and daughters among Swedes born between 1699 and 1871 (see Appendix Table A1). Marital sorting on socioeconomic status prevented extreme sex-patterning in socioeconomic transmission. Among Swedes born between 1785 and 1888, sons' occupational status was correlated with the occupational status of their fathers and, for daughters, the occupational status of their husbands was correlated with their fathers' occupational status. The correlation was somewhat weaker for daughters than for sons but differences were small (see Appendix Table A2).

1.2 Harvests in Sweden

We constructed annual harvest indices for each Swedish county (län), where we classified each harvest as either bad, normal, or good. (Details on construction of the harvest and other series we created are in the Appendix.) We used our harvest indices to classify a year as one of radical change (alternatively labeled a harvest swing) if either that year's harvest was good and the prior year's harvest was bad or if that year's harvest was bad and the prior year's harvest was good. Our year-county panel is not balanced prior to 1792, when we are more likely to have data for northern Sweden, where climactic conditions

were harsher.

Figure 1, which maps the fraction of extreme harvest swing years (years with a change from good to bad or bad to good harvests) by county, demonstrates that the fraction of harvest swings was greater from 1792 to 1839 than from 1840 to 1888. The figure also shows that northern relative to central and southern Sweden experienced more radical change in yearly harvests both before and after 1840.

Low agricultural surpluses exacerbated harvest shocks. The probability of a poor harvest following a bad harvest was higher prior to 1840 relative to after 1840 (see Appendix Table A3), suggesting that before 1840 seed grain was consumed in bad harvest years. While poor harvests are associated with elevated current mortality rates only prior to 1789 (see Appendix Table A4), food insecurity persisted and the years 1800/1801, 1808/1809, and 1867/1868 were famine years for some northern counties (Dribe et al., 2015).

Agricultural surpluses rose with the increase in average seed yields for grains after 1820 (see Table 6.1 in Morell et al. (2011)) when Sweden went from being a major importer of grains to an exporter. The change coincides with the policy switch to free trade and institutional reform to increase production incentives, including the abolition of interior tolls in 1810 and of export regulations in 1828, the sale of Crown land, and privatization of land from enclosures of 1803, 1807, and 1827. Canal building lowered transportation costs across counties. A railroad building boom, first private and then public, began in the 1840s and 1850s. Improved incentives led peasants to become agricultural entrepreneurs (Persson, 1999; Gadd, 2011).

Farmers learned to adapt to droughts, but not to unusually cold winters, after 1839. Drought increased the probability of a bad harvest between 1792 and 1839 but not between

1840-70 (see Appendix Table A5). Adaptation to droughts came from diversification between and within grain varieties sensitive to summer temperature and precipitation during different months (Skoglund, 2022).

Small agricultural producers arguably fared poorly both in lean and fat years prior to 1840, facing high prices for seed grain after a weak harvest and selling surpluses at low prices.[6] Low prices and high transport costs to more distant markets led small farmers to consume excess harvests rather than sell them. Saving excess harvests was risky because of destruction from damp weather, vermin, and fire, and discounted prices for last year's grain. The early policy response to poor harvests was to control prices to prevent hoarding and to establish state and parish storehouses (magasins).[7] A drawback to storehouses was uncertainty in how communally stored grain would be allocated. Central government supported storehouses were abolished by Parliament in 1824 and relief from bad harvests came from loans, in-kind aid, and public works projects which hired the unemployed (Persson, 1999).[8] The importance of parish storehouses peaks after 1830 for most of Sweden (Berg, 2007) with the exception of Sweden's northern-most counties where parish storehouses were set up after 1836 (Nelson, 1988).[9] These counties were not connected to the rest of the country by rail until the 1870s, leading to isolation when the Bering Sea was frozen over.

We find that own bad harvests reduced prices by at least 11 percent prior to 1840 (see

[6] This point was lamented by contemporaries (Persson, 1999).

[7] One of the aims of parish storehouses was to provide seed in case of crop failure.

[8] Parish storehouses continued but lost their importance by the 1870s (Berg, 2007).

[9] Parish storehouses were not necessarily successful in providing food in times of hunger because they were concerned with their own financial viability. Nelson (1988) describes the problems in getting aid to individuals at the local level in Norrbotten during the 1867-68 famine.

the first column of Appendix Table A6). After 1840, the relationship between prices and own poor harvests disappeared. Own good harvests were associated with roughly 6-7% lower grain prices in both time periods (see columns 1 and 3). Neighbors' bad harvests increased prices prior to but not after 1840 and neighbors good harvests lowered prices in both time periods (see columns 2 and 4), suggesting that there was economic integration with neighbors.

2 Data

Our analytical samples come from two sources. Our primary source consists of crowd-sourced genealogies on Family Search, augmented by linking to Swedish censuses. Our second source consists of multigenerational data in the SwedPop database on everyone who generated a record in selected parishes. We have linked the SwedPop data to additional death information, including for all of Sweden.

The Family Search sample consists of multigenerational data on 4,783 men and 4,453 women who lived to age 40, who were born before 1910 in an identifiable county in Sweden and who died in Sweden. The median birth year for both men and women was 1862. Everyone in the sample has four grandparents born before 1820 for whom we have harvest information when they were age 0-20. We linked grandparents at each age from 0 to 20 to our yearly county-level harvest index.[10] (See the Appendix for details on data construction and Appendix Figure A1 for a schematic of the sample.)

[10] We investigated the use of temperature and price data to predict harvest type and thus changes in harvest conditions but, because we explain very little of the variance, our predicted radical change series were uncorrelated with our measure of radical change derived from harvests. We therefore analyze only changes in harvest conditions.

The grandchildren in our analytical sample are descended from 4,673 fathers and 4,677 mothers who in turn were descended from over 4,200 fathers and mothers. The 4,276 paternal grandfathers had a median birth year of 1794. The records contain information on birthplace and, less often, on death place. We assume that an ancestor resided in a county at ages 0-20 if the ancestor was born in the county and either died in the county or his or her child was born in the county. We located grandparents in 1890 county boundaries and classified agricultural counties as those with a high agricultural and low manufacturing and mining value added in 1800.[11] We classified grandparents' parishes as agricultural if either the open soil sown was above 50% or if the parish had more than 75% farmers or cottagers.

The death year of the median grandson (born in 1862 and died in 1933) is the same as that from the yearly national cohort life table available in HMD. Human Mortality Database (2024). The median granddaughter (also born in 1862 and died in 1933) was shorter-lived by 2 years compared to the national cohort life table.

The northern-most and sparsely populated Swedish counties of Västerbotten and Norrbotten and the island of Gotland are over-represented in the data, with 47% of paternal grandfathers born in those counties (see Appendix Table A7.) An advantage of this over-representation is that much of our variation in harvest swings comes from northern Sweden. Robustness tests omit the counties of Västerbotten and Norrbotten.

We linked our analytical sample to the 1880, 1890 and 1900 censuses, which come from the extracts made by vicars and priests from parish books recording vital events.

[11]The non-agricultural counties were Stockholm, Blekinge, Göteborg och Bohus, Gävleborg, Örebro, Kopparberg, Värmland, Vastmanland, and Södermanland.

Seventy-five percent of both men and women were linked to at least one of the censuses. Census linkage revealed that our analytical sample is almost twice as agricultural compared to the population of Sweden as a whole. (See the Appendix for details.)

We created a socioeconomic index by letting a Lasso select census variables predictive of years lived between ages 40 and 80. Census variables included two-digit household head occupation, disability, family size, region of birth and of residence, and interactions between all variables. Our socioeconomic index is thus a summary measure of all potential mediating factors. We estimated four separate indices, by sex and farm status. Our estimation sample of 350,000 individuals was a randomly drawn 15% sample of the entire Swedish population born between 1840 and 1880, alive at age 40, and in at least one of the 1880-1900 censuses. (See the Appendix for details.) As an alternative measure of socioeconomic status we use the hiscam codes, where a higher code indicates greater socioeconomic status.[12]

We created an indicator variable for the grandchild experiencing a poor harvest in-utero and a continuous proxy for high contemporaneous infectious disease at every age between 40-80. Our indicator for contemporaneous infectious disease is the county-level infant mortality rate, which we compiled from both printed records and databases. (See the Appendix for details.)

We constructed samples of grandsons and granddaughters sharing the same great-grandfather (referred to as second cousin samples) to control for family effects. The grandsons and granddaughters in this sample are descended from grandfathers who were

[12] See https://www.camsis.stir.ac.uk/hiscam/. We used the universal scales for males (U2). A downside to these codes is that not all occupations are linked to hiscam codes.

brothers, where one grandfather experienced a harvest swing during the slow growing years prior to puberty and another one did not. Our first sample is from Family Search but we increased the window of grandparental birth year to births before 1850 and did not require information on all four grandparents. Compared to our main analytical sample the second cousin sample is more geographically skewed towards counties where the scope for adaptation remained limited, with 62% of paternal grandfathers born in Västerbotten or Norrbotten and 52% of paternal grandfathers experiencing harvest swings at ages 9-12. The Family Search second cousin sample contains 298 grandsons and 287 granddaughters,

Our other sample of second cousins is from family reconstitutions from parish records, available at SwedPop.[13] The original data are based on two different collections which cover a limited set of parishes within the counties of Västerbotten (the Demographic Data Base or DDB) and Malmöhus (the Scanian Economic-Demographic Database or SEDD). Although the public data do not follow mortality beyond 1912, we have linked grandsons and granddaughters to national death records from 1830 to 2020 (the Swedish Death Index), thus providing us both with death information on movers and on deaths beyond 1912. (The Appendix provides additional details on linkage and this sample.)

The grandsons and granddaughters in our second cousin SwedPop sample are descended from grandfathers who were brothers, where one grandfather experienced a harvest swing during the slow growing years prior to puberty and another one did not. This sample contains 1,941 grandsons and 2,002 granddaughters. The sample is restricted to grandchildren who survived to age 40 and known to be alive in 1831, when we have complete death record coverage. The second cousin SwedPop sample covers an earlier cohort

[13] See https://swedpop.se/.

than the Family Search sample: median birth and death years were 1833 and 1894 for grandsons and 1831 and 1892 for granddaughters. The second cousin SwedPop sample is more geographically skewed than the Family Search data, with almost 90% of paternal grandfathers born in Västerbotten because compared to the SEDD, the DDB is a larger sample, is more likely to contain three generations, and is more likely to cover generations linked to harvest data.

We created an additional sample of wives of the grandsons in our analytical sample drawn from FamilySearch to control for current family effects. The sample includes 3,192 wives of 3,191 grandsons who survived to age 40. (We allow for multiple marriages.)

There are different potential sources of bias in using crowd-sourced genealogies and parish-level family reconstitution data, in addition to geographic biases. First, in both data types, families with surviving descendants may be different from families without surviving descendants because of differences in both fertility and mortality. If families who were hit hardest by ancestral shocks have lower fertility and higher mortality, our estimates of the association between ancestral shocks and subsequent mortality may be biased downwards if we are missing the families most affected by harvest shocks or if fewer siblings leads to less competition for resources. Second, early childhood deaths are under-reported in genealogical data. Although chronic disease develops predominately at older ages, if the children prone to chronic disease die young we may be underestimating the strength of the association. In contrast, the family reconstitution will not under-report infant deaths. Third, the bias in family reconstitution data is that movers are lost to follow-up. We have minimized this bias among grandchildren by linking them to the Swedish Death Index, Fourth, the links across generations may be incorrect. Incorrect links are more likely in

large cities, areas which experienced extensive out-migration, and in incomplete family trees where there is less confirmation of family relationships.[14] We minimize the probability of incorrect links in the Family Search data by using family trees with information on all four grandparents. As a robustness test, we later expand the sample to include grandchildren who were born abroad or who emigrated, restricting to grandchildren with information on only the paternal grandfather but with information on the paternal great-grandfather.

3 Methods

If grandparents' harvest swings, arguably random events, are uncorrelated with individuals' genomes, their past environment, and their own past characteristics and the post-harvest characteristics of their grandparents then we can recover the causal relationship between longevity and harvest swings without controlling for these confounders. Our initial specification thus is

$$D_{i,t} = \beta_0 + \sum_G \sum_A \beta_{1,A,G} R_{A,G,i} + \beta_2 E_{i,t} + \beta_3 B_i + \beta_4 B_{G,i} + \beta_5 C_{G,i} \quad (1)$$

where $D_{i,t}$ equals 1 if the grandchild died in year t and is 0 otherwise and where we censor years lived to be between ages 40 and 80.[15] The variables $R_{A,G,i} = 1$ if there was radical change in the harvest quality of grandparent (G) at age group A and is 0 otherwise.

[14] Complete family trees are less likely to rely on on-line records and indices alone.

[15] Mortality at older ages is likely to have a different functional form. Our specification is equivalent to a Cox proportional hazard model. We present uncensored results and results for different initial age cutoffs in the Appendix.

The association between grandparents' harvest swings and grandchildren's mortality is effectively an interaction between grandparents' birth year and their county of residence and, because we cannot observe individual food availability, this association is an intent to treat. We estimate Equation 1 using a linear probability model where we control for fixed effects for own birth year (B_i) and for grandparents' birth years ($B_{G,i}$) and birth counties ($C_{G,i}$).[16] We cluster on the individual and run separate regressions for grandsons and granddaughters because of differences in mortality patterns by sex. The Appendix presents results with additional clustering on own and grandfathers' birth years and birth locations and with standard errors corrected for spatial correlation across grandchildrens' birth places.[17]

We categorize the slow growth period as ages 9-12 for boys and 8-10 for girls, thus assuming that the start of puberty was roughly 2 years later for the grandparents' generation compared to more recent generations.[18] Our full ancestral age categories are ages 0-2,3-5,5-8,9-12,13-16,and 17-19 for grandfathers and 0-2,3-4,5-7,8-10,11-13,14-16,and 17-19 for grandmothers. As a robustness test, we present single age estimates for paternal grandfathers' harvest swings. We do not take a stand on the association between controls for other ancestral ages and grandchildren's mortality, other than to argue that it should be

[16] For estimation see Correia (2017).

[17] To estimate the Conley (1999) standard errors we first demeaned using the code in Fetzer (2020) and then estimated the standard errors with a revised version of the code used in Hsiang (2010). We also estimated for every individual the probability of death between age 40-80, demeaned the data, and then ran synthetic outcome tests to construct new p-values (see Conley and Kelly (2025)).

[18] The slow growth period could be later in a deprived population. The age of menarche in Norway circa 1840 was 14 for upper class urban girls and 16 for lower class urban girls (Harland and Walløe, 1976). This would put the onset of puberty between ages 11-13 for girls and the slow growth period could range from ages 8-10 to 10-12.

smaller in less epigenetically sensitive time periods.[19]

Circumstantial evidence of epigenetic transmission requires that at least two conditions be met. The first is transmission of the paternal grandfather's harvest swings at ages 9-12 to grandsons but not to granddaughters. We use Benjamini-Hochberg tests to account for 52 hypotheses – 4 grandparents $\times$ 6 age categories for men plus 4 grandparents $\times$ 7 age categories for women.

We estimate Equation 1 for our full sample and for our sample restricted to agricultural counties and to agricultural counties and parishes. In ancestral areas more dependent on agriculture, with less diversified economies, the associations between ancestral harvest swings and grandchild mortality should be greater. We account for 102 hypotheses (our full ancestral age categories and agricultural vs. non-agricultural counties) using Benjamini-Hochberg tests.

The second condition for establishing transgenerational inheritance is that paternal grandfathers' harvest at ages 9-12 do not affect grandfathers' health and thus the health and socioeconomic status of their sons and their sons' sons. We will examine in the mechanisms section whether the whether harvest swings affect children's longevity, grandchildren's socioeconomic status, or fertility,

We modify Equation 1 to account for the contemporaneous infectious disease environ-

[19] Bygren (2013) reported that overfeeding of boys around age 3 increased the longevity of male-line grandsons whereas underfeeding of males about age 19 reduced the longevity of male-line grandsons.

ment, socioeconomic status, early life conditions, and parents' health,

$$D_{i,t} = \beta_0 + \sum_G \sum_A \beta_{1,A,G} R_{A,G,i} + \beta_2 E_{i,t} + \beta_3 B_i + \beta_4 B_{G,i} + \beta_4 C_{G,i} + \beta_5 \log(\text{IM}_{i,t}) + \beta_6 H_i + \beta_7 S_i + \beta_8 L_{F,i} + \beta_9 L_{M,i} + \beta_6 X_i \quad (2)$$

where $\log(\text{IM}_{i,t})$ is contemporaneous infant mortality by county, a proxy for the infectious disease environment, H_i is an indicator equal to one if the grandchild experienced a bad harvest in utero, S_i is a measure of the grandchild's socioeconomic status, and L_F and L_M are the lifespans of the father and of the mother. Additional controls (X_i) are latitude and longitude of the grandchild's birth parish and include fixed effects for the grandchild's death county. We also present specifications where we control for trends in county characteristics by including death county times birth year fixed effects.

The association between paternal grandfathers' harvest swings at ages 9-12 and grandsons' mortality should be stronger once we control for family effects. Only genes which are present can be switched on or off. Some families may be better able to adapt to harvest swings than others. We therefore use the second cousin samples where we have differential harvest swing exposure between grandfathers who were brothers. We estimate

$$D_{i,t} = \beta_0 + \beta_1 R_{FF,9-12,i} + \beta_2 \log(\text{IM}_{i,t}) + \beta_3 E_{i,t} + \beta_4 FFF_i + \beta_5 B_i \quad (3)$$

for grandchild i and year of life t between age 40 and 80 for the Family Search sample and

$$\text{Years Lived} = \beta_0 + \beta_1 R_{FF,9-12,i} + \beta_2 FFF_i + \beta_3 C_i \quad (4)$$

for the SwedPop sample. (We prefer to use lifepan for the SwedPop sample because the proportional hazards assumption is not met.[20]) The independent variables are $R_{FF,9-12,i} = 1$ if there was radical change in the harvest quality of the paternal grandfather when he was between age 9 and 12, $E_{i,t}$ are fixed effects for age, FFF_i is a fixed effect for the paternal great-grandfather, and B_i is a fixed effect for own birth year. We cluster on the father and on the grandfather because the samples includes siblings. For the Family Search sample we also cluster on the individual.

We modify Equation 1 to examine the association between grandsons' longevity and types of radical change (good to bad and bad to good harvest swings) and types of harvests (yearly radical change, at least one good and no bad harvests, at least one bad and no good harvests, and mixed but not radical change harvests). Examining types of radical change will help us understand whether variability per se or the direction of the variability matters. Bygren et al. (2000) find that in the Swedish county of Västerbotten in-utero exposure to any type of radical harvest change increased later life mortality. Investigating harvest types allows us to draw more comparisons with the Överkalix and Uppsala Multigen studies.

Our evidence for epigenetic transmission is a proof by elimination. We can rule in epigenetics and rule out alternative hypotheses, including socioeconomic factors, parental health, and differential fertility. Appendix Table A8 summarizes the predictions of each hypothesis. We examine if there are any direct effects of paternal grandfathers' harvest swings at ages 9-12 on their grandchildren's socioeconomic status and their children's mortality and fertility. We provide indirect evidence on the role of socioeconomic status by investigating the mortality of spouses, who should be affected by any family socioe-

[20]The Schoenfeld residuals for grandsons are correlated with time, χ^2 =5.85, p=0.054.

conomic factors. We establish if the evidence is consistent with mediation by the level of economic development by increasing the time window when grandparents were born (e.g., before 1850 rather than before 1820) because the association between ancestral harvest swings and grandchildren's mortality should fall over time with the increasing integration of markets and rising incomes.

4 Results

We first non-parametrically investigate the association between grandchildren's mortality and paternal grandfathers' harvest swings at ages 9 to 12. Kaplan-Meier survival curves, weighted by the inverse of the propensity score to account for differences in birth cohorts between the descendants of the "treated" and "untreated," show that radical change in harvests at the paternal grandfather's age of 9 to 12 is associated with lower survivorship among grandsons, particularly in agricultural counties and parishes.[21] Mortality differences between grandsons with exposed and unexposed paternal grandfathers are larger after age 60 and disappear by age 90. Although we find an association between paternal grandmothers' exposure to poor harvests at ages 8-10 for granddaughters, we also find an association for grandsons (see Figure 2). Because our results in Figure 2 could be affected by grandparent birth year and place effects, we next estimate Equation 1.

Figure 3 reveals that an extreme harvest swing for paternal grandfathers at ages 9-12 is associated with a statistically significant increase of 0.004 in grandsons' mortality at all ages between age 40 and 80 when we estimate Equation 1 with full age controls for all four

[21]The proportional hazards assumption is met: the Schoenfeld residuals for all grandsons are not correlated with time, χ^2 =0.13, p=0.719.

grandparents and with fixed effects to control for own and ancestral birth year and place effects. Our coefficient of 0.004 remains statistically significant at the 0.1 percent level, under alternative spatial error and clustering assumptions (see Appendix Table A9). We found no evidence of spatial dependence and estimated p-values were close to synthetic p-values (Conley and Kelly, 2025) (see Appendix Table A10). Our coefficient rises to 0.006 in agricultural counties and increases even further to 0.009 in agricultural parishes within agricultural counties. We find no correlation in non-agricultural counties areas. If grandsons' longevity and paternal grandfathers' harvest swings are correlated only at ages 40-80, our coefficients of 0.004, 0.006, and 0.009 imply respective decreases in life expectancy at age 40 within the sample of a half year, more than three-quarters of a year and more than a full year. There is no relationship between other paternal grandfathers' age groups and their grandsons' mortality. There also is no association between ancestral paternal grandfathers' harvest swings and granddaughters' mortality (see Figure 3).

There is at most a weak link between paternal grandmothers' harvest swings at ages 8-10 and granddaughters' mortality at ages 40-80 (see Appendix Figure A2) and no relationship at other ancestral ages. We find a coefficient of 0.003, which is statistically significant at the 5% level, when we restrict the sample to agricultural counties and find a small and statistically insignificant coefficient in the full sample.[22] Appendix Figures A3 and A4 plot the association of grandchildren's mortality with the harvest swings of maternal grandfathers and grandmothers.[23]

[22] Appendix Figure A2 shows that there was a statistically significant negative relationship between paternal grandmothers' extreme harvest swings at ages 17-19 and grandsons' mortality. We suspect that this is a chance relationship. The relationship is not statistically significant once we account for the number of hypotheses using Benjamini-Hochberg tests.

[23] Appendix Figure A4 shows that men whose maternal grandmothers experienced harvest swings at ages

The estimated relationship between grandsons' longevity and paternal grandfathers' harvest swings at ages 9-12, particularly in agricultural areas, is unlikely to arise from chance. Benjamini-Hochberg tests in Appendix Table A11 reveal that in the specifications run on grandsons with 26 grandparents' age groups, the p-values for paternal grandfathers' radical change at ages 9-12 always is below the critical value when the false discovery rate is set to 5%, with the exception of non-agricultural areas. If our hypothesis is that, consistent with transmission along the Y chromosome, paternal grandfather effects should be present for men and not for women, we have 52 hypotheses and our p-values for paternal grandfathers' radical change at ages 9-12 is always below the critical value when the false discovery rate is set to 10% and in agricultural areas is always below the critical value when the false discovery rate is set to 5%. In non-agricultural areas, the coefficient on grandpaternal harvest swings at ages 9-12 was not statistically significant. Comparing agricultural and non-agricultural areas gives us 104 hypotheses, and the p-value for the paternal grandfather's radical change at ages 9-12 is below the critical value when we set the false discovery rate to 5%. Associations of grandchildren's mortality with other grandparental harvest swings are not statistically significant once we account for 102 hypotheses using Benjamini-Hochberg tests.

The inclusion of control variables, including experiencing a poor harvest in utero, higher infectious disease (proxied by contemporaneous infant mortality) in each year of life between age 40 and 80, socioeconomic status and parents' longevity, does not affect the association between paternal grandfathers' radical harvest change at ages 9-12 (see Ta-

8-10 were less to die. However, this relationship is not statistically significant once we account for the number of hypotheses using Benjamini-Hochberg tests.

ble 1).[24] Those experiencing poor harvests in utero or higher contemporaneous infectious disease mortality and those with shorter-lived parents faced a higher risk of death in each year of life. Our socioeconomic index is highly statistically significant but the magnitude is modest. When we use hiscam codes as an alternative measure of socioeconomic status we find a modest but negative association between higher socioeconomic status and mortality, consistent with Dribe and Eriksson (2023).

Table 1 also reveals that the association between paternal grandfathers' radical harvest change at ages 9-12 and grandsons' mortality is of the same order of magnitude as the correlation between grandsons' mortality and a poor in-utero harvest. This association is greater than a half standard deviation increase in infant mortality in every year of life between age 40-80 or in our index of socioeconomic status.

Our estimated association between paternal grandfathers' harvest swings at ages 9-12 and grandsons' mortality rises from 0.004 to 0.010 when we control for original family effects by looking within families. Among male second cousins descended from paternal grandfathers who were brothers where one experienced a harvest swing at ages 9-12 and another did not, the risk of death at any age between ages 40-80 increased by 0.010 (see Table 2). This increase in the risk of death represents a loss of roughly 1.3 years of life at age 40. Among female second cousins, there was no relationship of mortality with paternal grandfathers' harvest swings. There was no link between mortality and paternal grandmothers' harvest swings at ages 8-10 for female second cousins whose maternal grandmothers were sisters where one experienced a harvest swing in that age group and

[24] We found small and statistically insignificant effects of a bad harvest in utero and of the logarithm of infant mortality rates on granddaughters' mortality. For granddaughters the coefficient on the index of socioeconomic status was -0.002 ($\hat{\sigma}$=0.000, p=0.000).

the other did not (results not shown).

We find a negative relationship between grandsons' years lived after age 40 and paternal grandfathers' harvest swings in the second cousin SwedPop dataset, with years lived falling by 1.4 if the paternal grandfather experienced a harvest swing at ages 9-12. (Using the logarithm of years lived as the dependent variable yields similar results.) There was no such relationship for granddaughters. In both the SwedPop and the Family Search samples, magnitudes were smaller and coefficients were statistically insignificant when we excluded the fixed effect for the paternal great-grandfather.[25]

All types of paternal grandfathers' radical change in harvest conditions at ages 9-12 are related to grandsons' mortality: experiencing swings from good to bad, bad to good, and both types of swings (see column 2 of Appendix Table A12). We did not find any statistically significant differences between types of harvest swings.

Grandsons whose paternal grandfathers experienced any type of yearly radical change at ages 9-12 faced a higher mortality than those whose paternal grandfathers experienced either a combination of normal harvests and at least one bad harvest, non-consecutive both good and bad harvests, or at least one good harvest and all normal harvests at ages 9-12 (see column 3 of Appendix Table A12). The good, bad and non-consecutive good and bad harvest categorizations were statistically indistinguishable from all normal harvests at paternal grandfathers' ages 9-12.

[25] When we use the full SwedPop sample, that is without the second cousin requirement, we only have statistical significance when we control for paternal great-grandfather fixed effects.

4.1 Robustness

The magnitude and statistical significance of the coefficient on paternal grandfathers' harvest swings at ages 9-12 does not change when we examine mortality at all ages above 39 (see the first column of Appendix Table A13). When we examine mortality at ages 40-60, 60-80, and 80 and older, we find that only grandsons' mortality at ages 60-80 was sensitive to paternal grandfathers' harvest swings at ages 9-12. The coefficient on paternal grandfathers' radical change was 0.008 (see the third column of Appendix Table A13), statistically significant at the 1% level and a statistically insignificant 0.001 at ages 40-60.[26] After age 80, mortality was strongly related to the infectious disease environment, as proxied by the yearly national infant mortality rate (results not shown). When we examine the probability of death prior to age one in the SwedPop second cousin sample, we find neither a statistically significant nor a meaningfully large relationship between the paternal grandfather's harvest swings and the probability of death (results not shown).

The association between grandsons' mortality and paternal grandfathers' radical change at ages 9-12 is not an artifact of how we defined the slow growth period. Appendix Figure A5 presents results examining ancestral radical change at single ages and grandchildrens' mortality. Paternal grandfathers' ages 11 and 12 were particularly important for transmission of harvest swings to grandsons. Paternal grandmothers' ages 9 and 10 were important for granddaughters in terms of magnitudes but the confidence intervals were wide.

Our results are not specific to our over-sampling northern Sweden. When we excluded

[26]The discontinuity is consistent with findings from a modern population showing a discontinuity for markers of metabolic disease around age 60 (Shen et al., 2024).

both Norrbotten and Västerbotten from the sample, the association between grandsons' mortality and paternal grandfathers' harvest swings at ages 9-12 in our main specification (Equation 1) was 0.005 ($\hat{\sigma}$=0.001, p=0.000).

Our findings are unlikely to be specific to our sample or to arise from emigration bias. Our results are noisier than but consistent with our prior findings when we expand the sample to almost 61,000 grandsons and more than 53,000 granddaughters by including grandchildren who either emigrated from Sweden or were born abroad, restricting to grandchildren with a known paternal great-grandfather and with information on the paternal grandfather's harvest swings at ages 9-12 but not necessarily at other ages or for other grandparents.[27] Appendix Table A14 shows that a paternal grandfather's radical harvest change reduced grandsons' years lived after age 40 by 0.9 years if the paternal grandfather was born anywhere in Sweden and by 1.5 years if the paternal grandfather was born in an agricultural county, controlling for fixed effects for birth year, grandfather's birth year and county and the paternal great-grandfather.[28] The association between paternal grandfathers' harvest swings and granddaughters mortality was not statistically significant.

5 Mechanisms

The epigentic hypothesis predicts not just sex-specific transmission with a critical period for transmission prior to puberty, but also increased likelihood of epigenetic inheritance when in-utero conditions are poor. Grandfathers' radical change is more closely associated

[27] Roughly 83% of the sample was born in Sweden.

[28] The relationship was small and statistically insignificant if we did not include a fixed effect for the paternal great-grandfather. Recall that in the SwedPop data our results also were sensitive to the controls for the paternal great-grandfather.

with grandsons' mortality if both grandsons and their fathers were born in the first half of the year, when food and nutrient availability was lower (see Appendix Table A15). Risk of death in each year was highest for grandsons of exposed paternal grandfathers born in the first semester with a father also born in the first semester. The second highest risk of death was for grandsons of exposed paternal grandfathers born in the second semester with fathers born in the first semester.[29]

Alternative explanations for sex-specific transmission include socioeconomic status, parental health, and differential fertility among fathers descended from exposed grandfathers.[30] We find no evidence that our results in Figure 3 are driven by socioeconomic status. Socioeconomic status was orthogonal to radical harvest change when paternal grandfathers were age 9 to 12. A regression of our socioeconomic index on paternal grandfathers' harvest swings at ages 9-12 yielded a small and statistically insignificant coefficient of 0.005 ($\hat{\sigma}$=0.050, p=0.925). If our paternal grandfather effect is a later life socioeconomic effect, it should predict the mortality of both husbands and wives because both share the same household but it does not (see Appendix Table A16). Chronic disease may develop at older ages when socioeconomic status already is established.[31]

We can rule out worse parental health driving our findings. There is no relationship between the mortality of fathers and mothers in our sample and their fathers' harvest swings (see Appendix Table A18), consistent with the prior literature finding evidence of gen-

[29] We did not find any statistically significant interactions between a grandfather's harvest swings at ages 9-12 and other variables, including socioeconomic status or a bad harvest in utero

[30] While periods of low food availability could also be accompanied by psychological stress, Black et al. (2016) find that the stress of the death of a grandparents while a child was in utero has no effect on the child's labor market or educational outcomes.

[31] A caveat is that we cannot measure productivity. Another caveat is that womens' mortality may be less sensitive than men's to socioeconomic status.

erational "skipping" (Kaati et al., 2007; Bygren et al., 2001; Vågerö et al., 2022, 2018). When we regressed the grandfather's life span on his harvest swings at ages 9-12 and controlled for his birth year and county fixed effects, the relationship was positive but statistically significant only at the 10% level ($\hat{\sigma}$=1.145, p=0.069).

We exclude differential paternal fertility as an explanation. Appendix Table A19 shows that there is neither a statistically significant nor an economically meaningful relationship between paternal grandfathers' exposure to harvest swings and the number of children, of either sex and age at death, born to the fathers in our sample. When we included the number of siblings as a control in Equation 1 we found that each additional sibling decreased grandsons' mortality rates by 0.0005 ($\hat{\sigma}$ =0.0002, p=0.003). As expected from the absence of a relationship between fertility and paternal grandfathers' exposure to harvest swings, there was no change in either the magnitude or the statistical significance of paternal grandfathers' harvest swings at age 9-12.

The link between grandsons' mortality and paternal grandfathers' harvest swings at ages 9-12 weakens as we include later cohorts in the sample, suggesting that the level of economic development mediates the impact of harvest swings. Our estimated coefficient of 0.004 in Figure 3 remains unchanged if we restrict grandparents' birth years to be prior to 1810 (see the first column of Appendix Table A20) but decreases if we include grandsons whose paternal grandfathers had birth years of 1820 or later (see columns 2-4 of Appendix Table A20).

The positive relationship between paternal grandfathers' exposure to harvest swings at ages 9-12 and their grandsons' older age mortality does not necessarily contradict prior transgenerational studies which found a negative association with over-feeding and a pos-

itive association with under-feeding (Kaati et al., 2007; Bygren et al., 2001; Vågerö et al., 2022, 2018; Van den Berg and Pinger, 2016). Over-feeding in a broad age range is ultimately a change between normal and good conditions because there are no cases of consistent over-feeding in a broad age range. Over-feeding, under-feeding, and radical change in feeding could all be epigenetic triggers. Homeostasis leads to a response at the cellular level and the cell remembers.

We cannot rule out differences in the external environment, including the level of economic development, as an explanation for differences between our study and prior transgenerational studies. Cohorts differed across studies and in the Uppsala studies none of the grandparents experienced mixed harvests. The relationship of paternal grandfathers' over-feeding to cause of death differed in the Överkalix and Uppsala studies[32], raising the possibility of interactions with later life environments.[33] The absence of a relationship between ancestral harvest exposures at other ages and grandchildrens' mortality contrasts with Bygren (2013) identifying age 19 as an epigenetically sensitive age and with Costa (2023, 2024) finding an association between grandsons' longevity and overweight and their paternal grandfathers' experiences of harsh captivity if they were age 19-24 when POWs. Outside of prison camps or isolated communities, young men could migrate at the first sign of trouble whereas in the typical household on the edge of subsistence young

[32] Kaati et al. (2002) and Pembrey et al. (2006) found a relationship with cardiovascular disease and diabetes for cohorts born in 1890 and 1905 (but not 1920) whose paternal grandparents experienced overfeeding. Vågerö et al. (2022, 2018) found an association between paternal grandparents' over-feeding at ages 9-12 with grandsons' cancer and cancer mortality in cohorts born in 1932 or later and alive in 1961 with ancestors born in 1874 or later.

[33] Luke et al. (2021) argue that in India it is the interaction between a poor ancestral environment in utero and rapid increases in adult nutrition with economic development that have led to the development of metabolic disease. Sekhri and Shastry (2024) find that diabetes rose in areas of India where high yield crops were introduced.

children would be the most vulnerable.

6 Explaining Mortality Trends

Table 1 showed that paternal grandfathers' harvest swings were roughly as detrimental to grandsons' longevity at older ages as high contemporaneous disease rates and experiencing a poor harvest in-utero. We use the regression coefficients in column 4 of Table 1 and national life tables for ten year male birth cohorts to investigate the implications for longevity of eliminating ancestral harvest swings and improving the contemporaneous disease environment. Because we do not have harvest data for all Swedish counties prior to 1792, we examine grandsons born between 1830 and 1909, using a limited set of Swedish counties for the 1830-1839 birth cohort. We estimate life expectancy under the no harvest swings counterfactual by assuming that in each year of life x between age 40 and 80 the mortality rate (m_x) fell by $0.005 \times f_{FF,9-12}$, where $f_{FF,9-12}$ is the national fraction of paternal grandfathers who experienced harvest swings between ages 9-12. We do not adjust mortality rates after age 80 when grandsons' mortality rates no longer were correlated with paternal grandfathers' harvest swings at ages 9-12 (see Appendix Table A13).

We examine the impact of improvements in the disease environment by assuming that the national infant mortality falls (or in some rare cases, rises) to the level in ten years' time. We interpret this "time-machine" scenario as a change in the technology. Under this scenario, mortality rates would fall by $0.005 \times \Delta \ln(m_x)$, where $\Delta \ln(m_x)$ is the change in the logarithm of mortality rates under each disease counterfactual.[34]

[34]We assume that the contemporaneous disease environment continues to influence mortality rates after

Table 3 reveals that ten year birth cohorts from 1830 to 1909 experienced first a rise and then a fall in the fraction with paternal grandfathers exposed to harvest swings at ages 9-12.[35] The fraction of paternal grandfathers exposed to harvest swings peaks at 0.5 for the 1860-69 cohort whose grandfathers were exposed between 1801 and 1813 and coincides with the poor harvests, world-wide mercantilist policies and trade disruptions of the French Revolutionary and Napoleonic Wars.[36] The fraction of paternal grandfathers exposed to harvest swings at ages 9-12 decreases precipitously for the 1880-89 birth cohort .

The elimination of harvest swings, as achieved by counties such as Malmöhus after 1801, would have increased life expectancy by a high of 1.37 years for the 1860 birth cohort and a low of 0.26 years for the 1900 birth cohort (see Table 3). Cohorts born before 1870 would have experienced the life expectancy of cohorts born roughly 20 years later. For comparison to risks today, men age 50-54 in 1992 lost two years of life if they were current smokers, a half year of life if they were obese and 0.24 years of life if they were high school rather than college educated (Østbye and Taylor, 2004).

Table 3 demonstrates that for men the elimination of harvest swings was as important as declines in contemporaneous infectious disease for the 1830-1870 birth cohorts.[37] Beginning with the 1880 birth cohort, using the technology available ten years later dom-

age 80.

[35] We assume that 68 years separate the generations and weight harvest swings by county population shares.

[36] See O'Rourke (2006) for an account of the worldwide economic impact of the French Revolutionary and Napoleonic Wars.

[37] A full accounting of infectious disease reductions should include multigenerational effects. We attempted to control for infectious disease outbreaks using episodes of mortality increases without accompanying rye price changes identified in Dribe et al. (2015) but these did not have any predictive power. We note that the evidence on in-utero exposure to the 1918 Influenza Pandemic is mixed (Almond, 2006; Beach et al., 2022; Wójcik et al., 2022; Cook et al., 2019).

inates the elimination of harvest swings with the largest gain of 1.47 years of life for the 1890 cohort.

Poor harvests in utero had a limited impact on the mortality of cohorts born between 1840 and 1910 because a relatively small proportion of the Swedish population experienced poor in-utero harvests after 1820. For example, only 8% of the 1840-1849 cohort experienced a poor harvest in-utero, implying that the elimination of poor harvests in utero would have led to an increase in life expectancy at age 40 of only 0.07 years. In contrast, for the 1800 birth cohort, 46% of whom experienced a poor harvest in utero, abolishing poor harvests in utero would have produced an increase in life expectancy at age 40 of 0.61 years.

7 Conclusion

Economic theories of consumption posit that individuals will smooth their consumption – the value of additional consumption declines with the level of consumption and, in the extreme case, low consumption leads to starvation. An additional and unexpected benefit of consumption smoothing in a poor and undeveloped economy is the increase in the longevity of grandchildren, particularly male-line grandsons.

What lessons can be drawn from how consumption smoothing was achieved in the past? Persson (1999) (p. 113) argued that market integration brought increased price stability to Europe and increased farmers' incomes by allowing them to take advantage of good harvests and thus smooth their consumption. Surpluses were more likely when farm implements improved and grains became less important when potatoes, a source of both

nutrition and diversification, were widely adopted (Gadd, 2011). Campbell and Ó Gráda (2011) emphasized changes in yield variability, showing that in England and Great Britain yield variability declined over the long run, with the coefficient of variation of gross wheat and barley yields falling by more than 40% between the time periods 1750-1850 and 1889-1939. Decreases in yield variability required innovations in selecting and sowing seeds more adaptable to different climate conditions. The introduction of autumn rye and autumn wheat in Sweden in the late nineteenth century both permitted a two to three field system without the need for a year of fallow after the preceeding harvest and provided diversification because the new varieties were sensitive to climate shocks during different summer months (Skoglund, 2022). In the United States, the adoption of new crop varieties led to an increase in the range over which corn could grow (Olmstead and Rhode, 2008). A precondition for adopting potentially risky innovations is the ability to smooth consumption between good and bad times, suggesting that market integration and crop innovation were self-enforcing once peasants were given the incentives to become agricultural entrepreneurs.

References

Almond, Douglas (2006) "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of Political Economy*, 114 (4), 672–712.

Almuwaqqat, Zakaria, Qin Hui, Chang Liu et al. (2024) "Long-Term Body Mass Index Variability and Adverse Cardiovascular Outomes," *JAMA Network Open*, 7 (3), 10.1001/jamanetworkopen.2024.3062.

Alsan, Marcella and Claudia Goldin (2019) "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880-1920," *Journal of Political Economy*, 127 (2), 586–638.

Anderson, Mark D., Kerwin Kofi Charles, and Daniel I. Rees (2022) "Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality," *American Economic Journal: Applied Economics*, 14 (2), 126–157.

Barker, DJP (1992) *Fetal and Infant Origins of Adult Disease*, London: British Medical Publishing Group.

——— (1994) *Mothers, Babies, and Disease in Later Life*, London: British Medical Publishing Group.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro (2016) "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century," *Journal of Political Economy*, 124 (1), 105–159.

Beach, Brian, Ryan Brown, Joseph Ferrie, Martin Saavedra, and Duncan Thomas (2022) "Reevaluating the Long-Term Impact of In Utero Exposure to the 1918 Influenza Pandemic," *Journal of Political Economy*, 130 (7), 1963–1990.

Berg, Bengt Åke (2007) *Volatility, Integration and Grain Banks. Studies in Harvests, Rye Prices and Institutional Development of the Parish* Magasins *in Sweden in the 18th and 19th Centuries* Ph.D. dissertation, Stockholm School of Economics, Available at: https://research.hhs.se/esploro/outputs/991001480459406056 [Accessed August 2024].

Van den Berg, GJ and PR Pinger (2016) "Transgenerational effects of childhood conditions on third generation health and education outcomes," *Economics and Human Biology*, 23 (Dec), 219–229.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2016) "Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes," *American Economic Journal: Applied Economics*, 8 (1), 193–223.

Bleakley, Hoyt and Sok Chul Hong (2017) "Adapting to the Weather: Lessons from U.S. History," *Journal of Economic History*, 77 (3), 756–9.

Bygren, Lars Olov (2013) "Intergenerational Health Responses to Adverse and Enriched Environments," *Annual Review of Public Health*, 34 (1), 49–60.

Bygren, Lars Olov, Sören Edvinsson, and Göran Brostrom (2000) "Change in Food Availability During Pregnancy: Is it Related to Adult Sudden Death From Cerebro- and Cardiovascular Disease in Offspring," *American Journal of Human Biology*, 12, 447–453.

Bygren, Lars Olov, G Kaati, and S Edvinsson (2001) "Longevity determined by paternal ancestors' nutrition during their slow growth period," *Acta Biotheoretica*, 49 (1), 53–9.

Bygren, Lars Olov, Petter Tinghöf, John Carstensen, Sören Edvinsson, Gunnar Kaati, Marcus Pembrey, and Michael Sjöström (2014) "Change in paternal grandmothers' early food supply influenced cardiovascular mortality of the female grandchildren," *BMC Genomic Data*, 15 (12).

Campbell, Bruce M.S. and Cormac Ó Gráda (2011) "Harvest Shortfalls, Grain Prices, and Famines in PreIndustrial England," *Journal of Economic History*, 71 (4), 858–886.

Chen, Qi, Wei Yan, and Enkui Duan (2016) "Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modification," *Nature Reviews*, 17 (December), 733–743.

Conley, Timothy G. (1999) "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92 (1), 1–45.

Conley, Timothy G. and Morgan Kelly (2025) "The standard errors of persistence," *Journal of International Economics*, 153.

Cook, C. Justin, Jason M. Fletcher, and Angela Forgues (2019) "Multigenerational Effects of Early-Life Health Shocks," *Demography*, 56 (5), 1855–1874.

Correia, Sergio (2017) "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator," http://scorreia.com/research/hdfe.pdf.

Costa, Dora L (2023) "Overweight grandsons and grandfathers' starvation exposure," *Journal of Health Economics*, 91 (September), 102796.

——— (2024) “Grandchildren’s Longevity and Their Grandfathers’ POW Trauma in the U.S. Civil War,” *Demography*, 61 (2), 337–361.

Currie, Janet (2011) “Inequality at Birth; Some Causes and Consequences,” *American Economic Review*, 101 (3), 1–22.

Cutler, David and Grant Miller (2005) “The role of public health improvements in health advances: The twentieth-century United States,” *Demography*, 42 (1), 1–22.

Deaton, Angus (2013) *The Great Escape; Health, Wealth, and the Origins of Inequality*, Princeton, NJ: Princeton University Press.

Dribe, Martin and Björn Eriksson (2023) “Socioeconomic Status and Adult Lifespan, 1881-2020. New Estimates from Swedish Death Registers and Full-count Census Data,” Lund Papers in Economic Demography, 2023:3.

Dribe, Martin, Mats Olsson, and Patrick Svensson (2015) “Famines in the Nordic Countries, AD 536-1875,” *Lund Papers in Economic History, Lund University, Department of Economic History* (138).

Easterlin, Richard A. (1996) *Growth Triumphant: The Twenty-first Century in Historical Perspective*, Ann Arbor, MI: University of Michigan Press.

Fernández-Villaverde, Jesús and Pablo A. Guerrón-Quintana (2020) “Uncertainty shocks and business cycle research,” *Review of Economic Dynamics*, 37, S118–S146.

Fetzer, Thiemo (2020) “Can Workfare Programs Moderate Violence? Evidence from India,” *Journal of the European Economic Association*, 18 (6), 3337–3375.

Floud, Roderick, Robert W. Fogel, Bernard Harris, and Sok Chul Hong (2011) *The Changing Body: Health, Nutrition, and Human Development in the Western World since 1700*, Cambridge: Cambridge University Press.

Forsdahl, A (1977) “Living conditions in childhood and subsequent risk factors for development of arteriosclerotic heart disease,” *Journal of Epidemiology and Community Health*, 32, 34–37.

Gadd, Carl-Johan (2011) “The Agrarian Revolution in Sweden,” in Myrdal, Janken and Mats Morell eds. *The Agrarian History of Sweden: 4000 BC to AD 2000*, 118–164, Lund, Sweden: Nordic Academic Press.

Harland, Gro and Lars Walløe (1976) “Menarcheal age in Norway in the 19th century: A re-evaluation of the historical sources,” *Annals of Human Biology*, 3 (4), 363–374.

Heijmans, BT, EW Tobi, AD Stein, H Putter, GJ Blauw, ES Susser, PE Slagboom, and LH Lumey (2008) "Persistent epigenetic differences associated with prenatal exposure to famine in humans," *Proceedings of the National Academy of Sciences, USA*, 104 (44), 17046–9.

HMD. Human Mortality Database (2024) "Max Planck Institute for Demographic Research (Germany) and University of California, Berkeley (USA), and French Institute for Demographic Studies (France)," www.mortality.org.

Hsiang, S.M. (2010) "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences*, 107 (35), 15367–72.

Jablonka, Eva and Gal Raz (2009) "Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution," *The Quarterly Review of Biology*, 84 (2), 131–176.

Kaati, G, LO Bygren, and S Edvinsson (2002) "Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period," *European Journal of Human Genetics*, 10, 682–688.

Kaati, Gunnar, Lars Olov Bygren, Marcus Pembrey, and Michael Sjöström (2007) "Transgenerational response to nutrition, early life circumstances and longevity," *European Journal of Human Genetics*, 15 (7), 784–90.

Kahn, Matthew E. (2010) *Climatopolis: How Our Cities Will Thrive in the Hotter Future*: New York: Basic Books.

Luke, Nancy, Kaivan Munshi, Anu Oomen, and Swapnil Singh (2021) "Economic Development, the Nutrition Trap, and Metabolic Disease," *National Bureau of Economic Research Working Paper Series* (29132).

Lumey, LH and FWA van Poppel (2013) "The Dutch Famine of 1944-45 as a Human Laboratory: Changes in the Early Life Environment and Adult Health," in Lumey, LH and Alexander Vaiserman eds. *Early Life Nutrition and Adult Health and Development*, 59–76, Hauppage, New York: Nova Science Publishers.

Morell, Mats, Carl-Johan Gadd, and Janken Myrdal (2011) "Statistical Appendix," in Myrdal, Janken and Mats Morell eds. *The Agrarian History of Sweden: 4000 BC to AD 2000*, 118–164, Lund, Sweden: Nordic Academic Press.

Murphy, KE, TG Jenkins, and DT Carrell (2016) "How the Father Might Epigenetically Program the Risk for Developmental Origins of Health and Disease Effects in His Offspring," in Rosenfeld, Cheryl ed. *The Epigenome and Developmental Origins of Health and Disease*, 361–75, London: Academic Press, Chapter 18.

Nätt, Daniel and Anita Öst (2020) "Male reproductive health and intergenerational metabolic responses from a small RNA perspective," *Journal of Internal Medicine*, 288, 305–320.

Nelson, Marie C. (1988) *Bitter Bread: The Famine in Norrbotten 1867-1868*, Acta Universitatis Upsaliensis. Studia historica Upsaliensia, 153, Uppsala, Sweden: [Uppsala University].

Olmstead, Alan L. and Paul W. Rhode (2008) *Creating Abundance: Biological Innovation and American Agricultural Development*, Cambridge, UK: Cambridge University Press.

O'Rourke, Kevin (2006) "The worldwide economic impact of the French Revolutionary and Napoleonic Wars, 1793-1815."

Østbye, Truls and Donald H Taylor (2004) "The effect of smoking on years of healthy life (YHL) lost among middle-aged and older Americans," *Health Services Research*, 39 (3), 531–52.

Özgüldez, Hatice Özge and Aydan Bulut-Karslioğlu (2024) "Dormancy, Quiescence, and Diapause: Savings Accounts for Life," *Annual Review of Cell and Developmental Biology*, 40, 16.1–16.25.

Pembrey, Marcus E., Lars Olov Bygren, Gunnar Kaati, Sören Edvinsson, Kate Northstone, Michael Sjöström, Jean Golding, and The Alspac Study Team (2006) "Sex-specific, male-line transgenerational responses in humans," *European Journal of Human Genetics*, 14, 159–166.

Persson, Karl Gunnar (1999) *Grain Markets in Europe, 1500-1900: Integration and Deregulation*, Cambridge, UK: Cambridge University Press.

Preston, Samuel H. (1975) "The Changing Relation Between Mortality and Level of Economic Development," *Population Studies*, 29 (2), 231–48.

Ray, Deepak, James S. Gerber, Graham K. MacDonald, and Paul C. West (2015) "Climate variation explains a third of global crop yield variability," *Nature Communications*, 6, 5989.

Reddy, Hemakumar, Rupa Bhattacharya, Shrish Tiwari et al. (2021) "Y chromosomal non-coding RNAs regulate autosomal gene expression via piRNAs in mouse testis," *BMC Biology*, 19 (198).

Ryan, Caleb P. and Christopher W. Kuzawa (2020) "Germline epigenetic inheritance: Challenges and opportunities for linking human paternal experience with offspring biology and health," *Evolutionary Anthropology*, 29, 180–200.

Samphantharak, Krislert and Robert M. Townsend (2018) "Risk and Return in Village Economies," *American Economic Journal: Microeconomics*, 10 (1), 1–40.

Sekhri, Sheetal and Gauri Kartini Shastry (2024) "The Curse of Plenty: The Green Revolution and the Rise in Chronic Disease," *American Economic Journal: Applied Economics*, 16 (4), 1–33.

Shen, Xiaotao, Chuchu Wang, Xin Zhou, Wendy Zhou, Daniel Hornburg, Si Wu, and Michael P. Snyder (2024) "Nonlinear dynamics of multi-omics profiles during human aging," *Nature Aging*.

Skoglund, Martin Karl (2022) "Climate variability and grain production in Scania, 1702-1911," *Climate of the Past*, 18 (3), 405–433.

Vågerö, Denny, Agneta Cederström, and Gerard J. van den Berg (2022) "Food abundance in men before puberty predicts a range of cancers in grandsons," *Nature Communications*, 13 (7507).

Vågerö, Denny, Pia R Pinger, Vanda Aronsson, and Gerard J. van den Berg (2018) "Paternal grandfather's access to food predicts all-cause and cancer mortality in grandsons," *Nature Communications*, 9 (5124).

Veenendaal, MV, RC Painter, SR de Rooij, PM Bossuyt, JA van der Post, PD Gluckman, MA Hanson, and TJ Roseboom (2013) "Transgenerational effects of prenatal exposure to the 1944-45 Dutch famine," *BJOG*, 120 (5), 548–53.

Wójcik, Juditha, Christian Bommer, Dominik Naeher, and Sebastian Vollmer (2022) "The long-term consequences of the global 1918 influenza pandemic: A systematic analysis of census data from 51 countries," *Discussion Paper No. 292, Georg-August-Universität Göttingen, Courant Research Centre – Poverty, Equity and Growth (CRC-PEG)*.

Yang, Zhen, Feng Xu, Andrew E Teschendorff, Yi Zhao, Lei Yao, Jian Li, and Yungang He (2022) "Insights into the role of long non-coding RNAs in DNA methylation mediated transcriptional regulation," *Frontiers in Molecular Biosciences*, 9, 106740.

Figure 1: Harvests Swings in Sweden

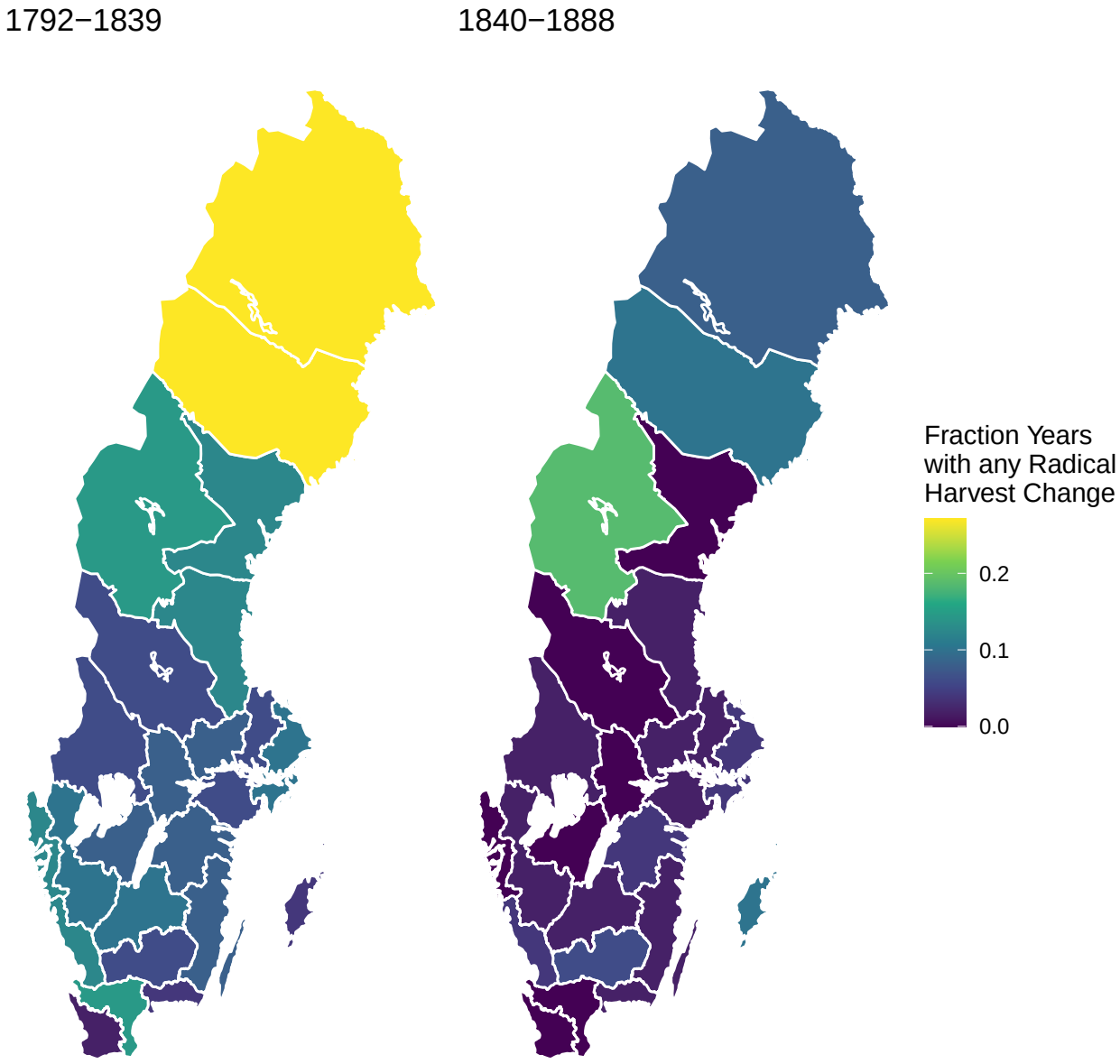
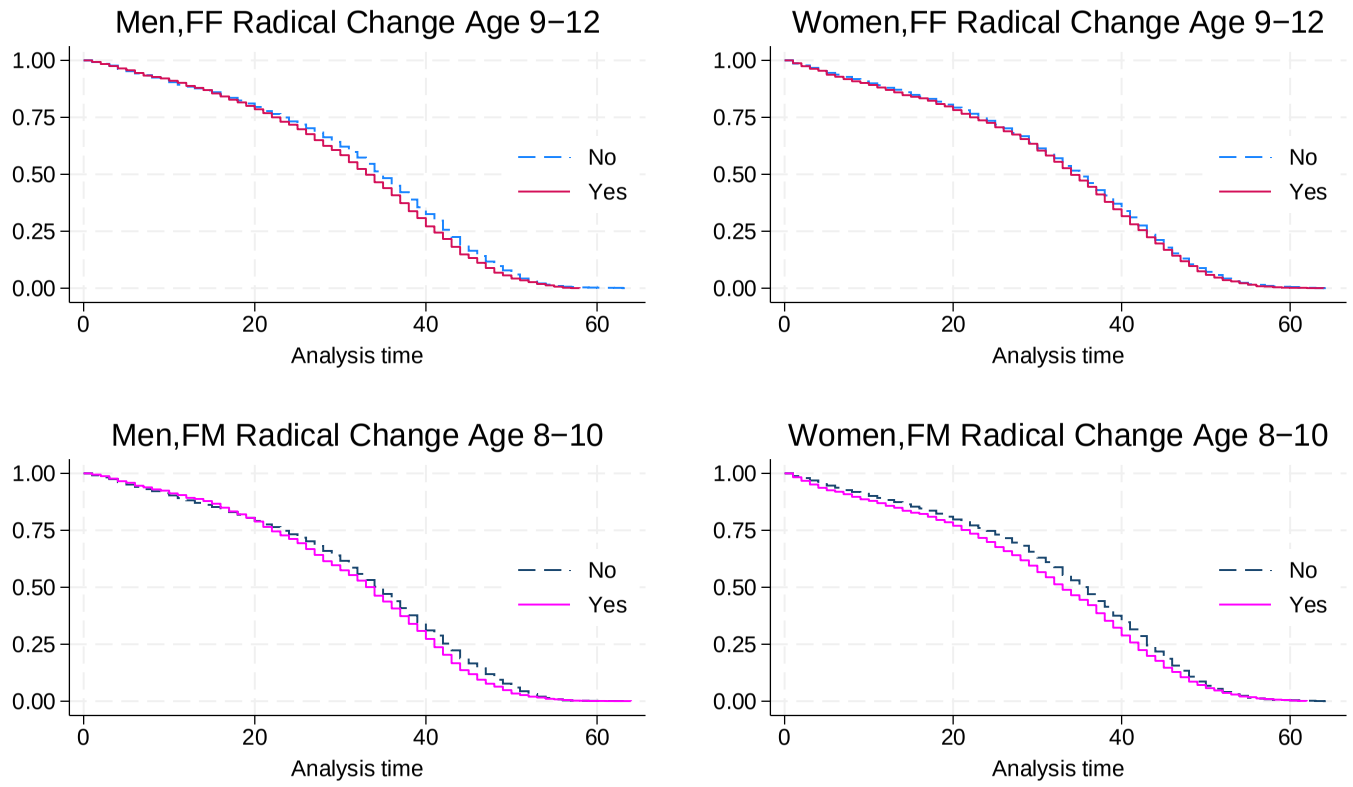
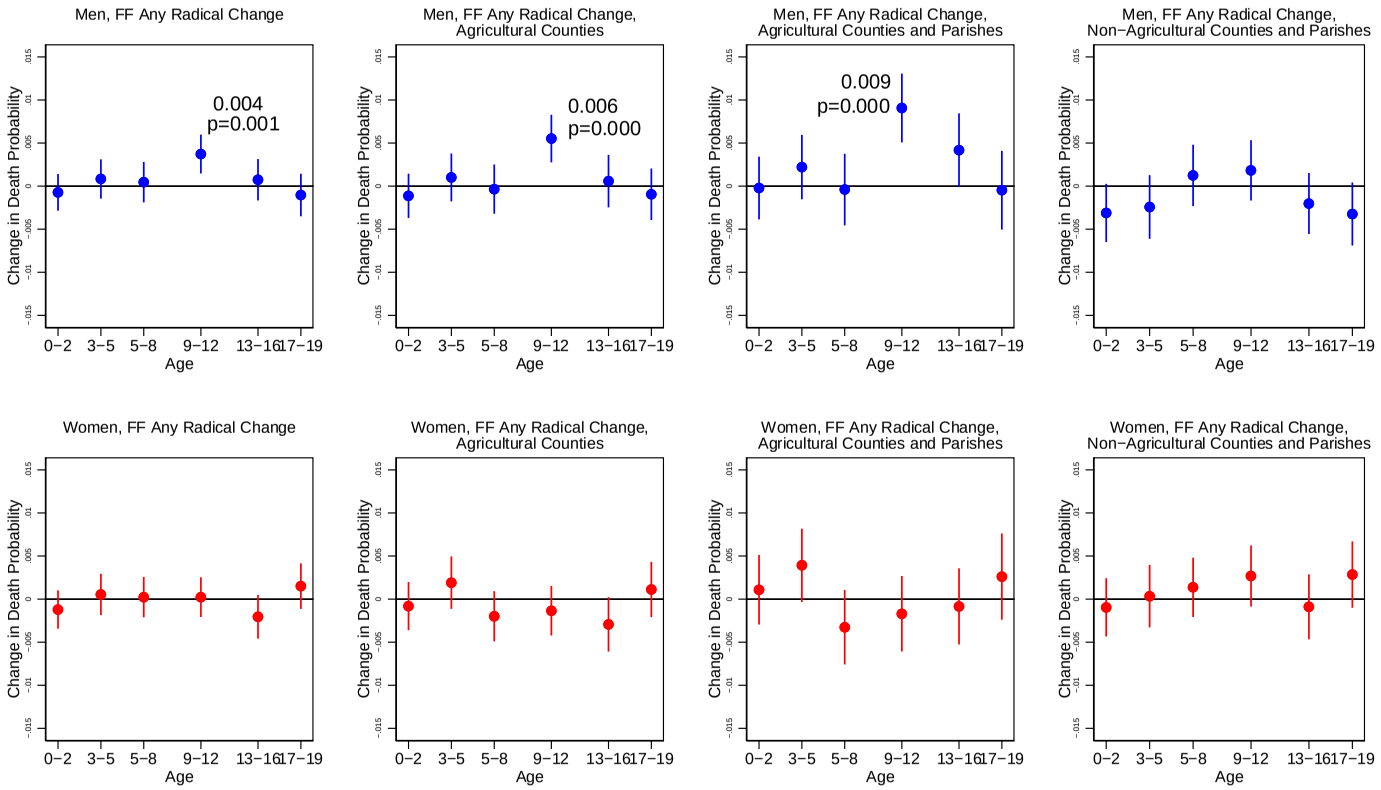


Figure 2: Grandsons' and Granddaughters' Kaplan-Meier Survival Curves and Paternal Grandfathers' Radical Change



The sample is weighted by the inverse of the propensity score which was estimated using single year of age and a common support. The estimate is the effect of the treatment on the treated. 4,471 grandsons and 4,386 granddaughters.

Figure 3: Grandsons' and Granddaughters' Mortality Increases at Ages 40-80 and Paternal Grandfathers' (FF) Radical Harvest Change



Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups. 95% confidence intervals are indicated by the lines. The number of observations for grandsons in each panel are 4,783 (all), 3,708 (agricultural counties), 1,969 (agricultural counties and parishes) and 2,814 (non-agricultural). The respective adjusted R^2 is 0.021, 0.021, 0.022 and 0.022. The number of observations for granddaughters in each panel are 4,451 (all), 3,416 (agricultural counties), 1,767 (agricultural counties and parishes) and 2,684 (non-agricultural). The respective adjusted R^2 is 0.019, 0.019, 0.023 and 0.017.

Table 1: Grandsons' Mortality and Paternal Grandfathers' Radical Change and Own Bad Harvest in Utero, Contemporaneous Disease, and Socioeconomic Status

Dependent variable: dummy=1 if died	(1)	(2)	(3)	(4)	(5)	(6)
Paternal grandfather experienced radical change	0.004***	0.004***	0.005***	0.005***	0.005**	0.007***
at ages 9-12	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
(mean for grandsons = 0.367)	[0.001]	[0.001]	[0.001]	[0.001]	[0.016]	[0.000]
Bad harvest in utero		0.004***	0.005**	0.004**	0.002	0.002
(mean for grandsons = 0.197)		(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
		[0.004]	[0.025]	[0.042]	[0.533]	[0.556]
Time-varying log(infant mortality rate)		0.006**	0.005**	0.005**	0.005	0.007**
(mean for grandson-years = -2.814)		(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
		[0.011]	[0.020]	[0.018]	[0.180]	[0.011]
Index of socioeconomic status				-0.002***		
(mean for grandsons = 1.706)				(0.000)		
				[0.000]		
Logarithm(hiscam)					0.016*	
(mean for hiscam grandsons = 51.716)					(0.009)	
					[0.016]	
(Father's Age at Death)/100						-0.012***
(mean for grandsons = 70.137)						(0.004)
						[0.001]
(Mother's Age at Death)/100						-0.009**
(mean for grandsons = 70.918)						(0.004)
						[0.022]
Effect of 1/2 std. dev. increase in						
Log(infant mortality rate)		0.002	0.001	0.001		
(Std. dev. = 0.543)						
Index of socioeconomic status				-0.001		
(Std. dev. = 1.460)						
Death county FE		Y	Y	Y		
Death county x birth year FE			Y	Y	Y	
Number of grandsons	4,783	4,783	4,782	4,782	2,259	3,644
Adjusted R^2	0.021	0.021	0.024	0.024	0.026	0.026

Estimated using Equation 2. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols *** and ** indicate significance at the 1 and 5% level, respectively. We were not able to code all occupations using the hiscam codes.

Table 2: Within Family Association Between Paternal Grandfathers' Radical Change and Grandchildren's Mortality

Sample:	Family Search		SwedPop			
Dependent variable:	Dummy=1 if died		Years Lived		Logarithm(Years Lived)	
	Grand-sons	Grand-daughters	Grand-sons	Grand-daughters	Grand-sons	Grand-daughters
Dummy=1 if the						
Paternal grandfathers experienced	0.010**	0.001	-1.423**	-0.591	-0.024**	-0.009
a harvest swing at ages 9-12	(0.004)	(0.004)	(0.597)	(0.589)	(0.010)	(0.009)
	[0.017]	[0.822]	[0.017]	[0.316]	[0.011]	[0.337]
logarithm(infant mortality rate)	0.007	0.007				
(time-varying by county and year)	(0.006)	(0.008)				
	[0.295]	[0.356]				
Number of individuals	298	287	1,941	2,002	1,941	2,002
Number of fathers	175	163	1,075	1,090	1,075	1,090
Number of paternal grandfathers	152	141	659	681	659	681
Adjusted R^2	0.03	0.03	0.146	0.159	0.141	0.154

Estimated using Equation 2. Standard errors, clustered on the individual and on the father, are in parentheses and p-values are in square brackets. The symbol ** indicates significance at the 5% level.

Table 3: Estimated Increase in Life Expectancy at Age 40 from the Elimination of Harvest Swings and the Decline in Infectious Disease Mortality

Birth Cohort	Male Life Expectancy at Age 40	Fraction Paternal Grandfathers With Radical Change at Age 9-12	Life Expectancy if No Radical Change	Life Expectancy Gain	Life Expectancy if at each age 40+ IMR is IMR in next 10 years	Life Expectancy Gain
1830-39*	29.8	0.297	30.52	**0.70**	30.14	**0.32**
1840-49**	30.5	0.298	31.21	**0.71**	30.81	**0.31**
1851-59	30.95	0.346	31.94	**0.99**	31.49	**0.54**
1860-69	31.55	0.463	32.92	**1.37**	32.22	**0.67**
1870-79	32.4	0.359	33.38	**0.98**	33.09	**0.69**
1880-89	33.21	0.145	33.66	**0.45**	34.32	**1.11**
1890-99	33.97	0.141	34.43	**0.46**	35.44	**1.47**
1900-09	34.89	0.076	35.15	**0.26**	36.27	**1.38**

* We use 20 out of 24 Swedish counties. For each county for which we have information on harvest swings for at least 5 years we calculated the probability of experiencing radical change in any four year period based on the estimated annual probability.
** We use all counties. For counties with some missing harvest years we imputed the fraction of harvest swings using the relationship with neighbors' fixed effects from a linear regression for the years for which we have complete data.
IMR = infant mortality rate. Male mortality at ages 40-80 is for all of Sweden and is from HMD. Human Mortality Database (2024).

Online Appendix

A.1 Data

A.1.1 County Level Data

We constructed our **harvest series** using the series in Hellstenius (1871); Utterström (1957); Statistika Centralbyrån (1865-1911, 1874-1910, 1949); Hallberg et al. (2017), as well as the occasional description from selected Mortalitetstabeller, Tabellverket (f.d. Tabellkommissionen), available on-line at the Riksarkivet, and from the Hushållningsjournaler, 1776-1813 available at https://ukforsk2.se/hushalldb/hushall.php. On the 7 point scale used by Hellstenius (1871) bad harvests were 0-2, normal harvests were 3-4 and good harvests were 5-6. Because other harvest series consisted of harvest descriptions or crop yield statistics, judgement calls were used to construct a series consistent across time and counties. For a discussion of the quality of different sources, see Edvinsson (2009).

We obtained **climate data** used in Lagerlöf, Nils-Petter (2010) from Nippe Lagerlöf. A caution in using the climate data is that temperatures and precipitation are imputed.

We constructed a **crop price index** using 1820 production weights from Statistika Centralbyrån (1949) and Jörberg (1972)'s price series for rye, barley, and oats. We deflated the crop price index using Edvinsson and Söderberg (2010). A drawback to the deflator is that crop prices are a large part of the cost of living index. Note that we have price data for the northern=most counties beginning only in 1811, leading our estimate of the relationship between bad harvests and prices to be biased downwards.

The national infant mortality rate is from HMD. Human Mortality Database (2024).

We compiled **county-level infant mortality rates** for every age between 40 and 80 as indicators of contemporaneous infectious disease. The sources we used were 1) SHiPS[38] aggregated to the county level for years prior to 1860; 2) for 1860-79, various years from *Bidrag till Sveriges officiella statistik. A, Befolknings statistik*[39]; 3) for 1880-1967, own calculations based on censuses and the Swedish Death Index (Sveriges Dödbok)[40]; and 4) for 1968 onward, the statistical database from Statistics Sweden, aggregated from municipality-level data.[41]

A.1.2 Family Search Data

We constructed our **multigenerational samples** by using the Family Search API to find all deceased individuals born in a Swedish county and their descendants. Our search for individuals who were born before 1910, who lived and died in Sweden, had a known birth and death year, and had information on all four grandparents and their birth counties yielded a sample of 60,650 men and women. When we further restricted the sample to those with grandparents where we had information on grandparents' childhood harvests and on own harvest the sample fell to 39,653. Because Family Search records for Sweden are predominately records of those born in the late 1800s and the early 1900s, less than half of these individuals had all four grandparents born before 1820. Further restricting the sample to those who survived to age 40 left us with a sample of less 9,000 individuals.

[38] https://www.umu.se/en/centre-for-demographic-and-ageing-research/infrastructure-at-cedar/open-data/ships/

[39] Available as on-line pdfs from Statistics Sweden at https://www.scb.se/

[40] Sveriges släktforskarförbund (2021). Sveriges dödbok 8: 1830-2020. (Version 8.0). Solna: Sveriges släktforskarförbund.

[41] Because the links on the Statistics Sweden web site, https://www.scb.se/, change frequently we do not provide full links.

The restriction to three generations leads to the over-representation of the northern-most counties counties of Västerbotten and Norrbotten.[42]

Graf (2021) describes the procedures to geolocate the data. We used Palm (2014)'s database to classify a parish as agricultural if either the open soil sown was above 50% or the parish had more than 75% farmers or cottagers.

We linked our multigenerational data to the Swedish **censuses** of 1880, 1890. and 1900. Sixty-seven percent of women and 65% of men alive in 1900 were linked to the 1900 census and 75% of both men and women alive in a census year were linked to either the 1880, 1890, or 1900 census (see Appendix Table A21).[43] Linkage rates to the Swedish censuses are high relative to US censuses because of the greater consistency of names and birth dates across sources.[44]

Census linkage reveals that our analytical sample is more agricultural than the population of Sweden as a whole. Among those age 20-55 in 1900, only 32% had no one in household working in agriculture and 45% had one household member working as a farmer (see Appendix Table A22). The comparable figures for Sweden overall were 51% and 26%. In our sample 5% of households had a member working as an agricultural laborer compared to 8% for Sweden overall.

We used a Lasso to select census variable to create a **socioeonomic index**. Our es-

[42]These counties were not big sources of LDS Church converts nor were they large sending counties to the US. Over-representation of these counties also is evident in the family trees found on myheritage.com. We suspect that over-represention is explained by the northern-most counties having richer genealogical sources.

[43]Linkage was to the on-line versions available as Minnesota Population Center (2019); The Swedish National Archives and Umeå University and the Minnesota Population Center (2014, 2011a,b).

[44]Swedish administrative units were smaller and parish priests had greater incentives to record information accurately compared to US census enumerators.

timation sample consisted of 350,000 individuals, a randomly drawn 15% sample of the entire Swedish population born between 1840 and 1880, alive at age 40, and in at least one of the 1880-1900 censuses. The outcome variable was years lived between ages 40 and 80. We estimated four separate indices, by sex and farm status using the unpenalized coefficients and divided the indices by 10. (Using the penalized coefficients made little difference.) The variables we use include two-digit household head occupation, region of birth and of residence, year of birth, other variables such as disability and family size, and interactions between all variables. The variance in explained mortality differed between our four groups, with the highest variance among non-farm men and the lowest among non-farm women (see Figure A6). Birth year and region explain much of the variance and occupation had the most explanatory power among non-farm men.

Our indicator variable for whether the grandchild experienced a **bad harvest in-utero** is based on harvest exposure in each month in-utero. We classify exposure to a bad harvest as exposure in any in-utero month. Harvests occurred from July to September. Prior to July, individuals would be consuming the prior year's harvest. In July, individuals would be consuming both the current and the prior year's harvest. In September, individuals would be consuming the current year's harvest. A grandchild conceived in September thus would be exposed to the harvest for that calendar year for the entire time period in-utero. A grandchild conceived in January would be exposed to two different harvests.

A.1.3 SwedPop Data

The **SwedPop data** we use consists of two different databases: the Demographic Data Base (DDB) and the Scanian Economic-Demographic Database (SEDD). DDB primarily

covers parishes in Västerbotten County's Skellefteåregion from 1680 to 1950 and is described in Westberg et al. (2016). (We do not have access to data which is less than 100 years old.) SEDD primarily covers five parishes in the county of Malmöhus from 1680 to 1812. The data are described in Dribe and Quaranta (2020). One drawback to the SEDD data is that information on harvest swings for Malmöhus is not available prior to 1792 and a second drawback is that Malmöhus experienced very little harvest variability. Because DDB is more than ten times larger than SEDD, covers a longer time period, and is more likely to include three generations, our analyses largely uses DDB data.

We linked the grandchildren in our three generation SwedPop sample to the national Swedish Death Index for 1830 to 2020 (Släktforskarförbund, 2023) and restricted the sample to those who survived to age 40. The death index data appear to be more complete beginning in 1831 rather than in 1830. We therefore restricted our analytical sample to grandchildren known to be alive in 1831 or later, thus creating a subsample where everyone is at risk of being linked to the death records.[45]

References

Bester, Alan C., Timothy G. Conley, and Christian Hansen (2011) "Inference with dependent data using cluster covariance estimators," *Journal of Econometrics*, 165 (2), 137–151.

Conley, Timothy G. and Morgan Kelly (2025) "The standard errors of persistence," *Journal of International Economics*, 153.

[45]Dropping this restriction does not materially change the results because with rare exceptions the grandchildren who were alive in 1831 or later were the ones who survived to age 40 and who had second cousins where the grandfather had a different harvest experience from their own grandfather.

Dribe, Martin and Luciana Quaranta (2020) "The Scanian Economic-Demographic Database (SEDD)," *Historical Life Course Studies*, 9, 158–172.

Edvinsson, Rodney (2009) "Swedish Harvests, 1665-1820: Early Modern Growth in the Periphery of the European Economy," *Scandinavian Economic History Review*, 57 (1), 2–25.

Edvinsson, Rodney and Johan Söderberg (2010) "The evolution of Swedish consumer prices 1290-2008," in Edvinsson, Rodney, Tor Jacobsson, and Daniel Waldenström eds. *Exchange Rates, Prices, and Wages, 1277-2008*, 412–452, Stockholm: Ekerlids Förlag and Sveriges Riksbank.

Graf, Benedikt A. (2021) "Geolocating Swedish Genealogical Records," https://drive.google.com/file/d/1AMa-JdnuWmN3OV7uj3zwEBLVjBF97QPo/view.

Hallberg, Erik, Lotta Leijonhufvud, Martin Linde, and Lennart Andersson Palm (2017) "Harvests in Sweden before the agrarian revolution (1.0) [Data set]. University of Gothenburg," https://doi.org/10.5878/002863.

Hellstenius, J (1871) "Skördarna i Sverige och deras verkningar," Statistik tidskrift, Stockholm: P,A, Norstedt and Söner.

HMD. Human Mortality Database (2024) "Max Planck Institute for Demographic Research (Germany) and University of California, Berkeley (USA), and French Institute for Demographic Studies (France)," www.mortality.org.

Ibramigov, Rustam and Ulrich K. Müller (2010) "T-statistic based correlation and heterogeneity robust inference," *Journal of Business and Economic Statistics*, 28 (4), 453–468.

Jörberg, Lennart (1972) *A History of Prices in Sweden 1732-1914. Volume 1. Sources, Methods, Tables*, Lund: CWF Gleerup.

Lagerlöf, Nils-Petter (2010) "Malthus in Sweden," Meeting Papers 790, Society for Economic Dynamics: https://ideas.repec.org/p/red/sed010/790.html.

Martinsson, Örjan (2002) "Tacitus. Population of Scandinavia," https://www.tacitus.nu/historical-atlas/population/scandinavia.htm [last accessed 8/29/2024].

Minnesota Population Center (2019) "Integrated Public Use Microdata Series. International. Version 7-2 [dataset]. Minneapolis. IPUMS.," http://doi.org/10.18128/D020.V7.2.

Palm, Andersson (2014) “The database Sweden 1570-1810: population, agriculture, land ownership - Agricultural statistics on parish level, 1810 (1.0) [Data set]. University of Gothenburg.,” https://doi.org/10.5878/002159.

Släktforskarförbund, Sveriges (2023) “Sveriges dödbok 1830-2020 [elektronisk resurs].”

Statistika Centralbyrån (1865-1911) “Hushållnings-Sällskapens Berättelser,” in *Bidrag till Sveriges Officiela Statistik. Jordbruk och Boskapsskötsel*, Stockholm: P,A, Norstedt and Söner.

——— (1874-1910) “Sammdrag af Kungl. Maj:ts Befallningshafvandes Årsväxtberättelser,” in *Bidrag till Sveriges Officiela Statistik. Jordbruk och Boskapsskötsel*, Stockholm: P,A, Norstedt and Söner.

——— (1949) “Minnesskrift med anledning av den Svenska Befolkningsstatistikens 200-Åriga Bestånd,” Statistika Meddelanden. Ser A. Band VI:4, Stockholm: K.L. Boktryckeri.

The Swedish National Archives and Umeå University and the Minnesota Population Center (2011a) “National Sample of the 1890 Census of Sweden, Version 2.0. Minneapolis: Minnesota Population Center [distributor].”

——— (2011b) “National Sample of the 1900 Census of Sweden, Version 2.0. Minneapolis: Minnesota Population Center [distributor].”

——— (2014) “National Sample of the 1880 Census of Sweden, Version 2.0. Minneapolis: Minnesota Population Center [distributor].”

Utterström (1957) *Jordbrukets arbetare. Levnadsvillkor och arbetsliv på landsbygden från friherstiden till mitten av 1800-talet.*, 2, Stockholm: Tryckeri AB Thule.

Westberg, Annika, Elisabeth Engberg, and Sören Edvinsson (2016) “A Unique Source for Innovative Longitudinal Research: The POPLINK Database,” *Historical Life Course Studies*, 3, 20–31.

Figure A1: Analytical Sample with Median Birth and Death Dates

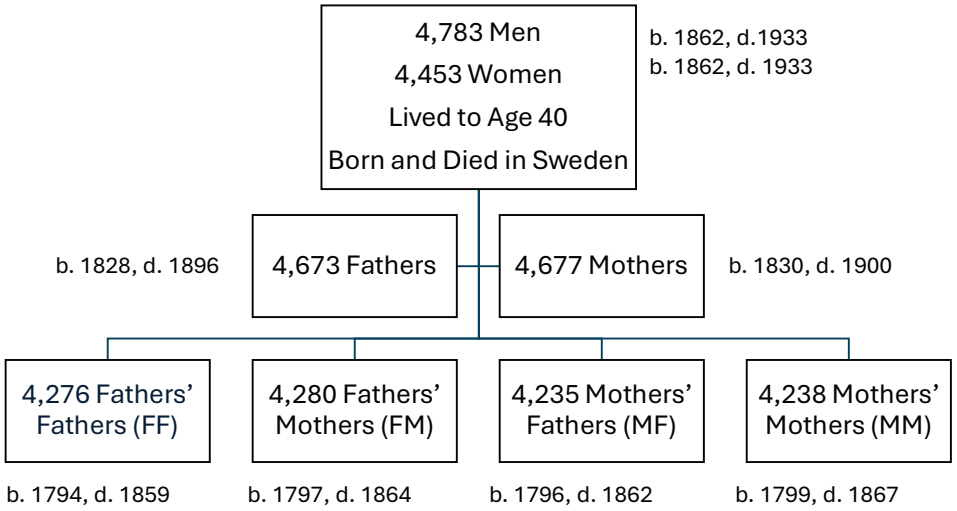
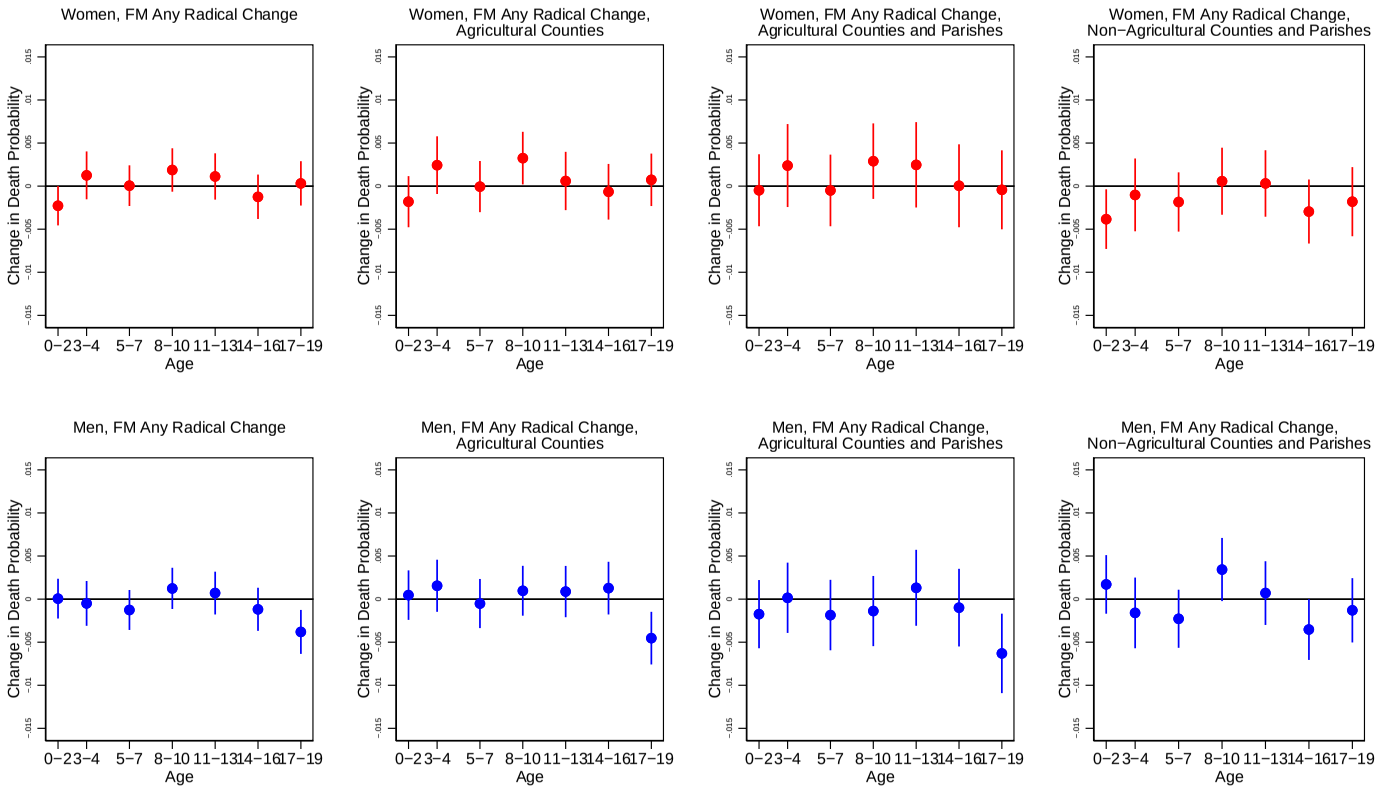
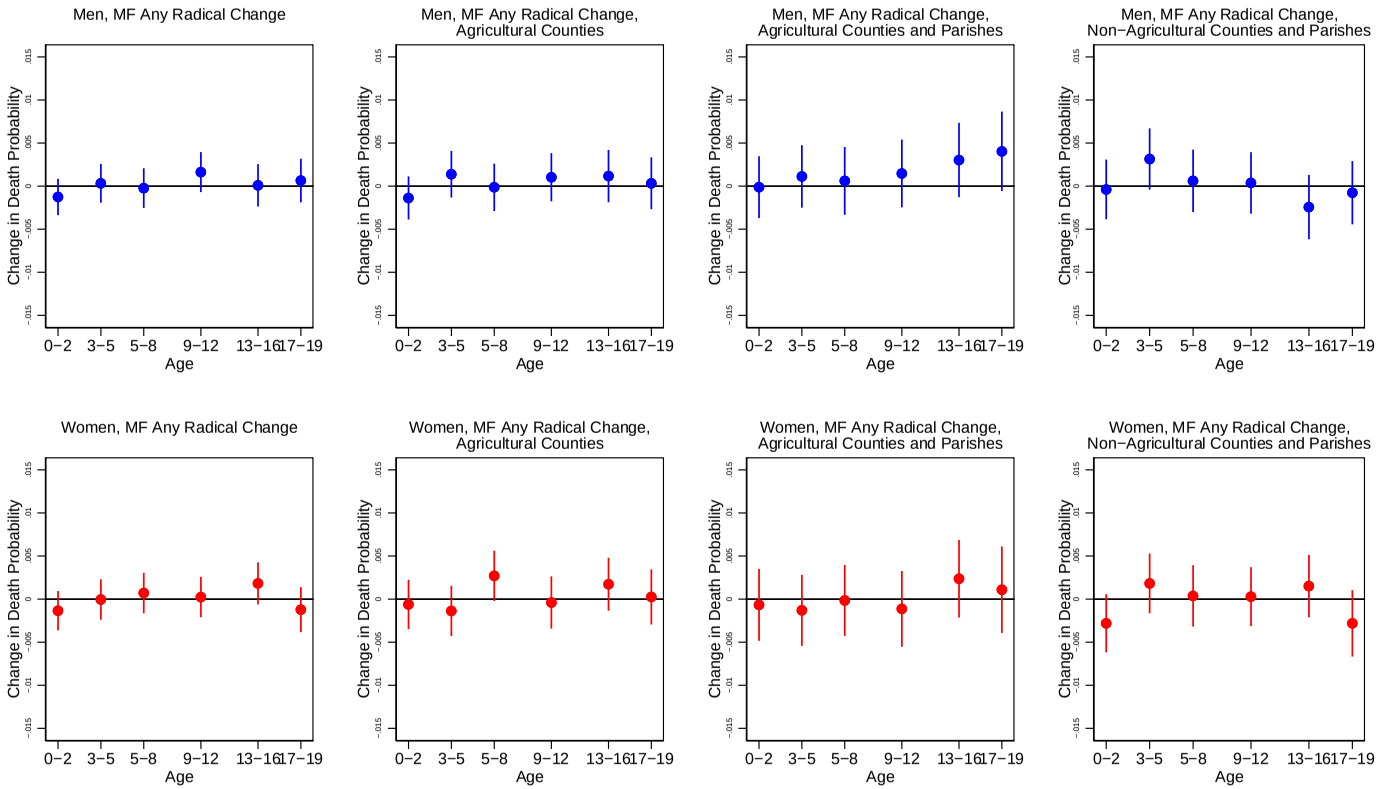


Figure A2: Granddaughters' and Grandsons' Mortality Increases at Ages 40-80 and Paternal Grandmothers' (FM) Radical Harvest Change



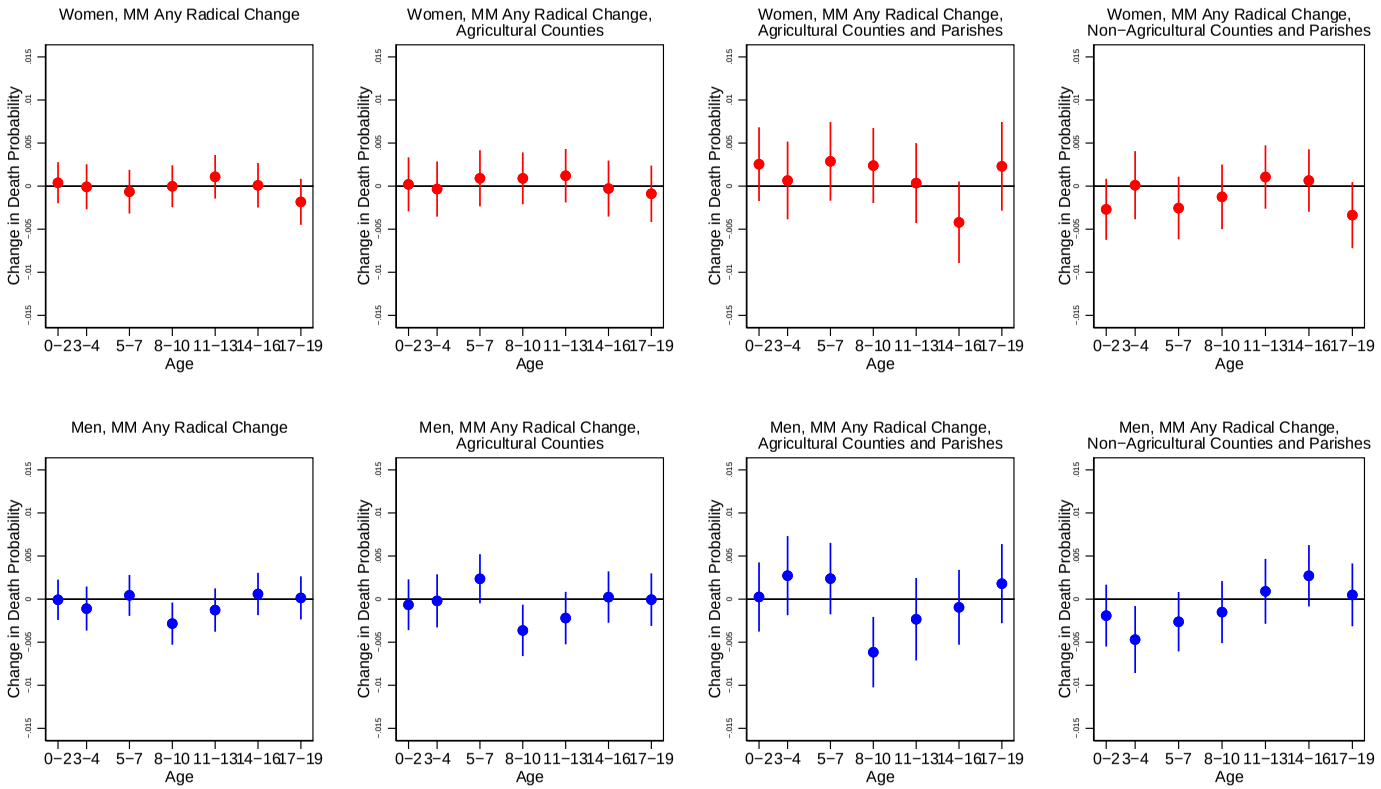
Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups. 95% confidence intervals are indicated by the lines. The number of observations in each panel for granddaughters are 4,451 (all), 3,421 (agricultural counties), 1,791 (agricultural counties and parishes) and 2,660 (non-agricultural). The respective adjusted R^2 is 0.019, 0.019, 0.021 and 0.18. The number of observations in each panel for grandsons are 4,783 (all), 3,714 (agricultural counties), 1,955 (agricultural counties and parishes) and 2,828 (non-agricultural). The respective adjusted R^2 is 0.021, 0.021, 0.021 and 0.023.

Figure A3: Granddaughters' and Grandsons' Mortality Increases at Ages 40-80 and Maternal Grandfathers' (MF) Radical Harvest Change



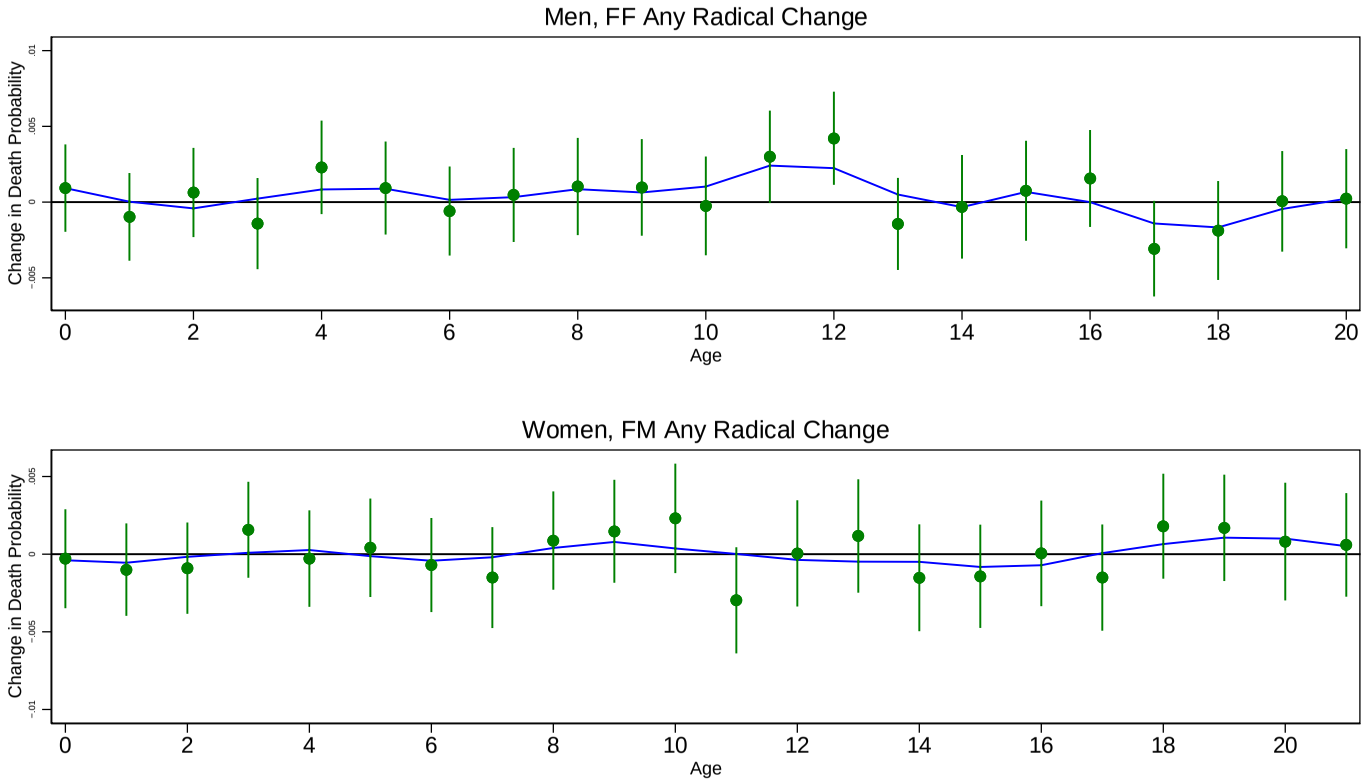
Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups. 95% confidence intervals are indicated by the lines. The number of observations for grandsons in each panel are 4,783 (all), 3,708 (agricultural counties), 1,969 (agricultural counties and parishes) and 2,814 (non-agricultural). The respective adjusted R^2 is 0.021, 0.021, 0.022 and 0.022. The number of observations for granddaughters in each panel are 4,451 (all), 3,416 (agricultural counties), 1,767 (agricultural counties and parishes) and 2,684 (non-agricultural). The respective adjusted R^2 is 0.019, 0.019, 0.023 and 0.017.

Figure A4: Granddaughters' and Grandsons' Mortality Increases at Ages 40-80 and Maternal Grandmothers' (MM) Radical Harvest Change



Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups. 95% confidence intervals are indicated by the lines. The number of observations for grandsons in each panel are 4,783 (all), 3,708 (agricultural counties), 1,969 (agricultural counties and parishes) and 2,814 (non-agricultural). The respective adjusted R^2 is 0.021, 0.021, 0.022 and 0.022. The number of observations for granddaughters in each panel are 4,451 (all), 3,416 (agricultural counties), 1,767 (agricultural counties and parishes) and 2,684 (non-agricultural). The respective adjusted R^2 is 0.019, 0.019, 0.023 and 0.017.

Figure A5: Mortality Increases at Ages 40-80 for 1) Grandsons and Yearly Radical Harvest Change of Paternal Grandfathers (FF) and 2) for Granddaughters and Yearly Radical Harvest Change of Paternal Grandmothers (FM)



Estimated using Equation 1 with radical change at each single year for paternal grandfathers only for grandsons and for maternal grandmothers only for granddaughters. 4,783 grandsons and 4,451 granddaughters. The adjusted R^2 is 0.021 for grandsons and 0.019 for granddaughters.

Figure A6: Variance Decomposition of Socioeconomic Index

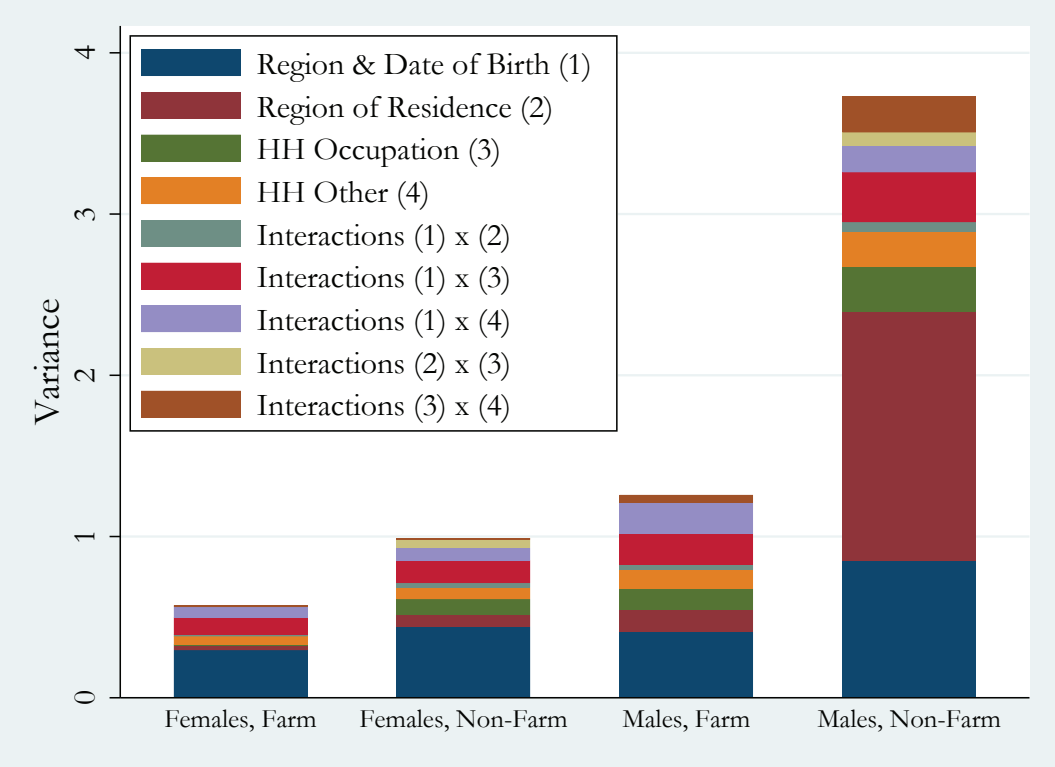


Table A1: Lifespan Correlations, Children and Parents

Dependent Variable: Years Lived	All Ages		Age 40+	
Years Lived	Men	Women	Men	Women
Variables:				
Father's life span	0.091***	0.104***	0.033***	0.034***
	(0.010)	(0.010)	(0.006)	(0.006)
	[0.000]	[0.000]	[0.000]	[0.000]
Mother's life span	0.157***	0.169***	0.039***	0.043***
	(0.009)	(0.009)	(0.006)	(0.006)
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	50,125	47,371	21,889	22,038
Adjusted R^2	0.209	0.229	0.193	0.216

Estimated using family reconstitutions from parish records, available at SwedPop (https://swedpop.se/), which have been linked to the deathbooks. We used OLS with a constant, fixed effects for birth year and birth county and with standard errors clustered on the father.

Table A2: Socioeconomic Status Correlations, Children and Fathers at Different Ages

Dependent Variable:	All			Västerbotten		
Son or Son-In-Law's Hiscam	Age 20	Age 40	Age 60	Age 20	Age 40	Age 60
Variables						
Dummy=1 if woman	2.532**	2.668***	-1.637	1.772	1.585	-3.461*
	(1.008)	(0.760)	(1.343)	(1.437)	(1.086)	(2.038)
	[0.012]	[0.000]	[0.223]	[0.218]	[0.144]	[0.090]
Father's hiscam at same age	0.296***	0.208***	0.151***	0.199***	0.117***	0.064***
	(0.016)	(0.012)	(0.017)	(0.027)	(0.018)	(0.021)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]
Woman $\times$ father's hiscam	-0.049**	-0.051***	0.037	-0.033	-0.029	0.076*
	(0.020)	(0.015)	(0.026)	(0.029)	(0.021)	(0.040)
	[0.016]	[0.001]	[0.159]	[0.259]	[0.172]	[0.056]
Constant	35.203***	40.566***	43.822***	40.165***	45.303***	48.271***
	(0.802)	(0.628)	(0.884)	(1.319)	(0.916)	(1.070)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	97,954	60,237	16,196	35,687	19,960	4,992
Adjusted R^2	0.091	0.057	0.047	0.047	0.032	0.038

Estimated using family reconstitutions from parish records, available at SwedPop (https://swedpop.se/). Socioeconomic status for both children and fathers is measured using hiscam codes. Husbands' hiscam is used for daughters. We used OLS with fixed effects for birth year and clustered the standard errors on the father. The first three columns also include fixed effects for county of birth.

Table A3: Probability of Good and Bad Harvests, Before and After 1840

	1793-1840	1840-1887
Prob(Good Harvest)	0.319	0.323
Prob(Good Harvest | Good Harvest Prior Year)	0.398	0.322
Prob(Good Harvest | Bad Harvest Prior Year)	0.274	0.385
Prob(Bad Harvest)	0.217	0.066
Prob(Bad Harvest | Bad Harvest Prior Year)	0.386	0.064
Prob(Bad Harvest | Good Harvest Prior Year)	0.132	0.069

Estimated using our constructed harvest index.

Table A4: Bad Harvests and Parish Mortality Rates

	1749-1788	1789-1819	1820-1859	
Dependent variable: log(crude mortality rate)				
Bad harvest	0.044***	0.002	0.008	
	(0.007)	(0.005)	(0.006)	
	[0.000]	[0.626]	[0.172]	
Difference (1749-1788) and (1789-1819)				0.042***
				(0.009)
				[0.000]
Difference (1749-1788) and (1820-1859)				0.036***
				(0.010)
				[0.000]
Difference (1789-1819) and (1820-1859)				-0.006
				(0.008)
				[0.432]
Observations	167,753			
Adjusted R^2	0.161			

Annual parish level crude mortality rates are from SHiPS,
https://www.umu.se/en/centre-for-demographic-and-ageing-research/infrastructure-at-cedar/open-data/ships/
For each time period, we estimated the specification $\ln(m_{it}) = D_i \times B_{it} + D_i \times C_i + T_t$, where m_{it} is the crude mortality rate in parish i at time t, B_{it} is a dummy equal to one if the harvest was bad in year i in parish t and is measured at the county level, C_i is a fixed effect for the county, and T_t is a fixed effect for the year. D_i is a fixed effect for the year group is not separately identified.

Table A5: Bad Harvests and Weather Shocks, Before and After 1840

	Average Marginal Effects		
Dependent Variable: Dummy =1 if Bad Harvest	1749-1839	1840-1870	Δ
Below average spring temperature	0.191***	0.172***	-0.018
	(0.025)	(0.042)	(0.049)
	[0.000]	[0.000]	[0.709]
Below average summer precipitation	0.105***	0.000	-0.104**
	(0.021)	(0.026)	(0.051)
	[0.000]	[0.983]	[0.040]

We thank Nippe Lagerlöf for providing the temperature and rainfall data. County smoothed temperature and precipitation were constructed using a Hodrick-Prescott filter on the logarithmic values. Below a logarithmic value of 0.15 is below average. Marginal effects are estimates from a probit equation where we allow for a random effect for the county. Our specification is $\Pr(H_{it} = 1|x_{it}) = \Phi(D_t + T_{it} + R_{it} + (D_t \times T_{it}) + (D_t \times R_{it}) + \nu_i)$, where H_{it} is equal to 1 if there was a bad harvest in county i at time t, D_t is equal to one if the year is after 1840, T_{it} is equal to one if spring temperature was below average, R_{it} is equal to one if summer rainfall was below average and ν_i is a random effect. Standard errors, clustered on the county, are in parentheses and p-values are in square brackets.

Table A6: Harvests and Harvest Prices

	Before 1840		After 1839		Δ	
	1	2	3	4	1-2	3-4
Dependent variable: log(crop price index)						
Bad Harvest	0.112***	0.068***	-0.001	-0.032*	0.112***	0.100***
	(0.011)	(0.013)	(0.017)	(0.019)	(0.021)	(0.023)
	[0.000]	[0.000]	[0.963]	[0.085]	[0.000]	[0.000]
Good Harvest	-0.077***	-0.034***	-0.066***	-0.035***	-0.010	0.001
	(0.010)	(0.011)	(0.008)	(0.009)	(0.013)	(0.014)
	[0.000]	[0.002]	[0.000]	[0.000]	[0.452]	[0.966]
Neighbor has a Bad Harvest		0.089***		0.017		0.072***
		(0.011)		(0.013)		(0.018)
		[0.000]		[0.198]		[0.000]
Good Harvest		-0.063***		-0.059***		0.001
		(0.010)		(0.009)		(0.014)
		[0.000]		[0.000]		[0.796]
Within R^2	0.079	0.128				
Between R^2	0.130	0.252				
Overall R^2	0.100	0.142				
Observations	3,190	3,190				
Number of counties	25	25				

We constructed the price index using 1820 production weights and Jörberg (1972)'s price series for rye, barley, and oats. The panel is unbalanced because we do not have price data for the northern-most counties prior to 1811. Our estimated specifications are $\ln(p_{it}) = D_t + B_{it} + G_{it} + C_i + D_t \times (B_{it} + G_{it})$ and $\ln(p_{it}) = D_t + B_{it} + G_{it} + NB_{it} + NG_{it} + C_i + D_t \times (B_{it} + G_{it} + NB_{it} + NG_{it})$, where p_{it} is our price index for county i at time t, D_t is a dummy variable equal to if the year was after 1839, B_{it} and G_{it} are dummies equal to one if the harvests were bad or good, respectively, for county i at time t, and NB_{it} and NG_{it} are dummies equal to one if the neighboring county had a bad or good harvest, respectively, and C_i is a fixed effect for the county. We allow for AR(1) disturbances in the error term.

Table A7: Geographical Distibution of the Analytical Sample and the Population of Sweden

	Percent Grandchildren Dying in County	Percent Paternal Grandfathers Born in County	All of Sweden 1930	1795
Stockholm (including city)	3.83	0.20	8.18	7.56
Uppsala	0.88	1.23	2.25	3.62
Södermanland	4.76	5.92	3.08	4.23
Östergötland	2.52	3.16	5.00	6.90
Jönköping	3.86	4.92	3.78	5.57
Kronoberg	4.03	5.39	2.54	3.81
Kalmar	4.30	4.19	3.76	5.66
Gotland	13.07	14.40	0.93	1.43
Blekinge	1.14	1.36	2.36	2.66
Kristianstad	1.22	1.13	4.01	5.04
Malmöhus	1.06	0.37	8.32	5.95
Halland	0.82	0.96	2.44	3.24
Göteburg och Bohus	1.40	0.93	7.44	4.47
Älvsborg	0.47	0.13	5.10	6.47
Skaraborg	0.28	0.20	3.94	5.95
Värmland	2.88	3.76	4.40	5.19
Örebro	2.52	3.06	3.57	4.19
Västmanland	1.41	0.83	2.64	3.71
Kopparberg	4.76	4.86	4.07	5.33
Gävleborg	2.05	1.80	4.56	3.14
Västernorrland	7.62	6.19	4.54	2.24
Jämtland	0.73	0.30	2.20	1.28
Västerbotten	19.46	19.06	3.32	1.09
Norrbotten	14.84	13.51	3.26	1.28
Unknown	0.11	0	0	0

Population numbers by county for Sweden are from Martinsson (2002).

Table A8: Predictions of Different Hypotheses

Mechanism	Prediction
Epigenetic transmission along the Y chromosome	Increased mortality of only grandsons if paternal grandfather was exposed to harvest swings at ages 9-12 Association stronger once control for family effects Association stronger if poor in-utero conditions No predictions about longevity of father or paternal grandfather
Socioeconomic status	Strong male-line relative to female-line socioeconomic transmission Paternal grandfather's exposure associated with grandsons' socioeconomic status Increased mortality of fathers born to exposed grandfathers Increased mortality among wives of male-line grandsons No predictions about age of paternal grandfather's exposure to harvest swings
Fertility	Differential fertility effects for exposed paternal grandfathers (mechanism could be either socioeconomic or biological through the mother)
Parental health	Reduced longevity among children born to exposed paternal grandfathers (mechanism could be either biological or socioeconomic) Reduced longevity among women marrying into the family (mechanism could be marriage markets)

Table A9: Estimates of Standard Errors on Paternal Grandfathers' Radical Change at Ages 9-12 Under Different Clustering and Spatial Standard Error Assumptions

	OLS, Double Clustering				Conely Standard Errors		
Dummy=1 if paternal grandfather	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
experienced radical change at age	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
9-12	[0.002]	[0.002]	[0.002]	[0.006]	[0.002]	[0.003]	[0.003]
Cluster also on							
own birth year	Y						
father's birth year		Y					
own birth parish			Y				
paternal grandfather's birth parish				Y			
Kilometers					300	50	20

Estimated using Equation 1. 4,783 grandsons. Kilometers are based on longitude and latitude of the birth location. The flight distance from Malmöhus in southern-most Sweden to Stockholm in central Sweden is roughly 500 kilometers. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbol *** indicates significance at the 1% level.

Table A10: P-values for Synthetic Outcome Tests for Paternal Grandfathers' Radical Change at Ages 9-12

	BCH		IM	
	Est.	Synthetic	Est.	Synthetic
Clusters	p	p	p	p
0	0.00	0.00	0.00	0.00
3	0.10	0.12	0.12	0.11
4	0.07	0.09	0.10	0.10
5	0.07	0.08	0.05	0.04
6	0.02	0.02	0.03	0.02
7	0.03	0.02	0.04	0.03
8	0.02	0.02	0.04	0.03
9	0.03	0.04	0.03	0.03
10	0.03	0.03	0.02	0.01

Our regression specification is Death Prob $= \beta_0 + \sum_G \sum_A \beta_{1,A,G} R_{A,G,i} + \beta_2 E_{i,t} + \beta_3 C_i + \beta_4 C_{G,i}$ where Death Prob is the probability that the individual died between ages 40-80, $R_{A,G,i} = 1$ if there was radical change in the harvest quality of grandparent (G) at age group A and is 0 otherwise; $E_{i,t}$ is a fixed effect for the age of the grandchild at time t; C_i is a fixed effects for own birth year and, $C_{G,i}$ are fixed effects for grandparents' birth years and birth counties. Our estimate of the coefficient on the paternal grandfather's radical change at ages 9-12 was 0.056. 4,692 grandsons (excludes grandsons where longitude and latitude of place of birth are unknown). The Moran I statistic was -0.38. The R^2 of a regression of the outcome on a quadratic in longitude and latitude was equal to zero. The synthetic p-values allow us to test whether we can reject the null hypothesis that the outcome is spatial noise. For estimation of synthetic p-values, see Conley and Kelly (2025) and the code at https://github.com/morganwkelly/spatInfer. We used 5 splines, one principal component, and 1,000 simulations. Because we find no evidence of spatial correlation in residuals, our preferred estimates are uncorrected for large clusters. However, we still present the Bester-Conley-Hansen (BCH) standard errors (Bester et al., 2011) and the Ibramigov and Müller (IM) standard errors (Ibramigov and Müller, 2010) with different numbers of clusters.

Table A11: Benjamini-Hochberg Critical Values for Effect of Paternal Grandfather's Radical Change at Ages 9-12 on Grandsons' Longevity

	Paternal Grandfather Radical Change at Age 9-12				
	p-value	p-value rank	FDR=.10	FDR=.05	Number of hypotheses
All grandfather birth places					
Grandsons	0.001209	1	0.003846	0.001923	26
Grandsons and granddaughters	0.001209	1	0.001923	0.000962	52
Grandfather born in agricultural county					
Grandsons	0.000090	1	0.003846	0.001923	26
Grandsons and granddaughters	0.000090	1	0.001923	0.000962	52
Grandfather born in agricultural county and parish					
Grandsons	0.000009	1	0.003846	0.001923	26
Grandsons and granddaughters	0.000009	1	0.000962	0.000481	52
Grandfather born in non-agricultural place					
Grandsons and granddaughters	0.306451	24	0.046154	0.023077	52
Grandsons and granddaughters, agricultural county and parish and non-agricultural places	0.000009	1	0.000962	0.000481	104

P-values are from Equation 1 (as estimated for Figures 3). The coefficient on paternal grandfathers' harvest swings at ages 9-12 in non-agricultural places (either a non-agricultural county or a non-agricultural parish in an agricultural county) was 0.002 ($\hat{\sigma}$ =0.002).

Table A12: Type of Radical Harvest Change of Paternal Grandfathers at Ages 9-12 and Grandsons' Mortality

	(1)	(2)	(3)
Paternal grandfather experienced radical			
change (36.7%) at ages 9-12	0.003***		
	(0.001)		
	[0.001]		
Type of radical change			
None (omitted, 63.3%)			
Good to bad (12.2%)		0.004**	
		(0.002)	
		[0.018]	
Bad to good (12.9%)		0.003*	
		(0.002)	
		[0.099]	
Both (11.6%)		0.003**	
		(0.002)	
		[0.049]	
Type of harvest			
Radical change (omitted,36.7%)			
Combination normal and at least one good (28.4%)			-0.004***
			(0.001)
			[0.008]
Normal harvest only (4.2%)			-0.002
			(0.002)
			[0.462]
Combination normal and at least one bad (21.8%)			-0.002*
			(0.001)
			[0.094]
Mixed harvest but no radical change (8.9%)			-0.003**
			(0.002)
			[0.093]
Adjusted R^2	0.021	0.021	0.021

Estimated using a variant of Equation 1 with controls for only the paternal grandfather's harvest characteristics at all age groups to avoid collinearity. 4,783 grandsons. Mean values in parentheses were calculated for all grandsons. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols *, **, and *** indicate significance at the 10, 5 and 1% level, respectively.

Table A13: Grandsons' Mortality and Paternal Grandfathers' Radical Change at Ages 9-12, Different Age Groups

Dependent variable=1 if died	No Censoring	Age 40-60	Age 60-80	Age 80+
Dummy=1 if paternal grandfather	0.004***	0.001	0.008***	-0.010
experienced radical change at	(0.001)	(0.001)	(0.002)	(0.013)
ages 9-12	[0.002]	[0.186]	[0.001]	[0.471]
Number of grandsons	4,783	4,783	3,784	1,436
Adjusted R^2	0.054	0.003	0.016	0.045

Estimated using Equation 1 and all four grandparents' radical change at all age groups. The last two columns of coefficients are conditional on survival to age 60 and 80, respectively. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbol *** indicates significance at the 1% level.

Table A14: Grandchildren's Years Lived after Age 40 and Paternal Grandfathers' Radical Change at Ages 9-12, Expanded Sample

	All		Agricultural County	
	Men	Women	Men	Women
Dependent variable: Years lived after age 40				
Dummy=1 if paternal grandfather	-0.884**	-0.635	-1.456***	-0.766
experienced harvest swing at	(0.419)	(0.493)	(0.513)	(0.593)
ages 9-12	[0.035]	[0.198]	[0.005]	[0.196]
Observations	60,952	53,522	42,498	37,081
R^2	0.346	0.365	0.343	0.361

The expanded sample includes all grandchildren born either in Sweden or abroad, with a known paternal great-grandfather and with information on paternal grandfather's harvest swings at ages 9-12 but not necessarily at other ages or for other grandparents Estimated from an OLS regression of years lived after age 40 on the paternal grandfathers' radical change at ages 9-12 with fixed effects for own birth year, grandfather's birth year and county, and the paternal great-grandfather. Standard errors, clustered on the grandfather, are in parentheses and p-values are in square brackets. The symbols ** and *** indicates significance at the 5 and 1% level, respectively.

Table A15: Own and Father's Semester of Birth and the Paternal Grandfather's Radical Change at Ages 9-12

	Paternal Grandfather Experienced Radical Change at Ages 9-12 vs None					
Dependent variable=1 if died	Coeffi-cient	Std. Err.	P-Value	Diff-erence With a)	Std. Err.	P-Value
Born in 1st semester						
a) and Father born in 1st semester	0.008***	0.002	0.000			
b) and Father born in 2nd semester	0.002	0.002	0.369	0.007***	0.003	0.009
Born in 2nd semester						
c) and Father born in 1st semester	0.004**	0.002	0.042	0.004*	0.002	0.073
d) and Father born in 2nd semester	0.000	0.002	0.865	0.008***	0.003	0.001
Adjusted $R^2 = 0.021$						

All coefficients are relative to no radical change, given own and father's semester of birth. Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups and modified to include an interaction between paternal grandfathers' radical change at ages 9-12 and own and father's semester of birth. 4,610 observations. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols *, **, and *** indicate significance at the 10, 5 and 1% level, respectively.

Table A16: Mortality Among Spouses

Dependent variable=1 if died	Husbands in Wife Sample	Wives in Wife Sample
Dummy=1 if paternal grandfather experienced radical change at ages 9-12	0.003** (0.001) [0.023]	
Dummy= 1 if husband's paternal grandfather experienced radical change at ages 9-12		0.001 (0.001) [0.503]
Adjusted R^2	0.023	0.017
Number of individuals	3,191	3,192

Estimated with controls for paternal grandfathers' radical change at different age groups. The number of wives is greater than the number of husbands because of remarriage. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols ** and *** indicate significance at the 5 and 1% level, respectively.

Table A17: The Association Between Fathers' and Childrens' Socioeconomic Status

	Circa Age 20		Circa Age 40		Circa Age 60	
	Men	Women	Men	Women	Men	Women
Dependent variable:						
Own or husband's hiscam						
Father's hiscam	0.298***	0.242***	0.209***	0.151***	0.142***	0.194***
at same age	(0.018)	(0.021)	(0.014)	(0.012)	(0.020)	(0.027)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Difference Men and	-0.056**		-0.058***		0.052*	
Women	(0.023)		(0.016)		(0.031)	
	[0.014]		[0.000]		[0.096]	
Observations	88,288		50,685		12,018	
Adjusted R^2	0.085		0.055		0.044	

Estimated using the SwedPop data (DDB and SEDD). The hiscam for women is that of their husbands. Estimated from a regression of own or husband's hiscam on a dummy indicating if female, own or husband's hiscam, the interaction between own and husband's hiscam, regional dummies and fixed effects for age. Clustered on the father. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols *,** and *** indicate significance at the 10, 5 and 1% level, respectively.

Table A18: Grandchildren's Parents' Mortality and Paternal Grandfathers' Radical Change at Ages 9-12

	Survival after birth first child in sample		Survival between Age 40-80	
	Grandchildren's			
	Fathers	Mothers	Fathers	Mothers
Dummy=1 if father experienced	0.001	0.000	-0.001	-0.001
radical change at ages 9-12	(0.001)	(0.001)	(0.002)	(0.002)
	[0.471]	[0.847]	[0.771]	[0.578]
Adjusted R^2	0.054	0.055	0.108	0.115
Number of fathers or mothers	3,706	3,567	2,637	2,364

Estimated using a modified form of Equation 1 where the dependent variable is whether the grandchildren's parents survived for every year of life. Every year of life is from the age of birth of the first child in the analytical sample until death in the first two columns and between age 40 and 80 in the last two columns. Controls include parents' parents' harvest swings at all age groups and included fixed effects for parents' own age, birth year, and death county (including unknown) and for parents' parents' birth year and birth county. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbol *** indicates significance at the 1% level.

Table A19: Fertility of Fathers and their Fathers' Harvest Swings at Age 9-12

	Dependent Variable		
	Number of Children	Number of Sons	Number of Daughters
Dummy=1 if grandchildrens' paternal grandfather			
experienced radical change at ages 9-12	0.015	0.036	-0.007
	(0.019)	(0.026)	(0.027)
	[0.413]	[0.164]	[0.798]
Number when no radical change	3.001	1.548	1.462
	(0.033)	(0.024)	(0.023)
	[0.000]	[0.000]	[0.000]
Number when radical change	3.057	1.605	1.452
	(0.044)	(0.032)	(0.031)
	[0.000]	[0.000]	[0.000]
Difference in number	0.047	0.057	-0.010
	(0.057)	(0.041)	(0.040)
	[0.414]	[0.166]	[0.798]
Pseudo R^2	0.032	0.024	0.025

Coefficients are from a Poisson regression where the dependent variable is the total number of children for the fathers in the sample. Controls are included for the fathers' fathers' harvest swings at all ages groups and fixed effects for fathers' birth year and birth county. 4,649 fathers. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets.

Table A20: Coefficient on Paternal Grandfathers' Radical Change at Ages 9-12 and Grandsons' Mortality Under Different Ancestral Birth Year Cut-offs

	All Ancestors born before			
	1810	1830	1840	1850
Dummy=1 if radical change	0.004**	0.002**	0.002**	0.001
when paternal grandfather	(0.002)	(0.001)	(0.001)	(0.001)
was ages 9-12	[0.010]	[0.020]	[0.023]	[0.123]
Adjusted R^2	0.021	0.021	0.022	0.023
Number of grandsons	3,289	6,997	9,381	11,040

Estimated using Equation 1 with controls for all four grandparents' radical change at all age groups. Standard errors, clustered on the individual, are in parentheses and p-values are in square brackets. The symbols *, **, and *** indicate significance at the 10, 5 and 1% level, respectively.

Table A21: Census Linkage Rates

	Full Sample			At Risk Linkage		
	All	Men	Women	All	Men	Women
1880	54.85	53.67	56.12	60.93	59.92	62.00
1890	62.10	61.72	62.5	65.98	65.77	66.21
1900	65.72	64.46	67.09	73.21	72.32	74.15
Any Census	74.78	74.60	74.97	76.23	76.11	76.36

Those at risk of linkage were those alive during the census years. Linkage was to the on-line versions of the Swedish census available as Minnesota Population Center (2019) and
The Swedish National Archives and Umeå University and the Minnesota Population Center (2014, 2011a,b).

Table A22: Household Participation in Agriculture

	Age 20-55 in 1900	
	Analytical Sample	General Population
No one in the household works in agriculture	33.59	51.10
At least one household member works as a farmer	44.80	25.8
At least one household member works as an agricultural laborer	5.63	7.45
At least one household member works as a cottar*	9.23	7.45
Unknown	6.76	8.2

*A cottar occupies a cottage and small holding of land in return for services.
Percentages for the population sample were calculated from Minnesota Population Center (2019) and The Swedish National Archives and Umeå University and the Minnesota Population Center (2014, 2011a,b).