Socioeconomic Status and Adult Lifespan, 1881-2020. New Estimates from Swedish Death Registers and Full-count Census Data

MARTIN DRIBE, BJÖRN ERIKSSON

LUND PAPERS IN ECONOMIC DEMOGRAPHY 2023:3

Centre for Economic Demography
Lund University School of Economics and Management
P.O. Box 7083
SE-220 07 Lund, Sweden

www.ed.lu/publications/Lund-Papers-in-Economic-Demography/

Socioeconomic Status and Adult Lifespan, 1881-2020.
New Estimates from Swedish Death Registers and Full-count Census Data

Martin Dribe and Björn Eriksson

Centre for Economic Demography
Department of Economic History
Lund University
P.O. Box 7080
220 07 Lund, Sweden

Martin.Dribe@ekh.lu.se
Bjorn.Eriksson@ekh.lu.se

Acknowledgements

This study is part of the research program "The Rise and Fall of the Industrial City: Landskrona Population Study", funded by the Swedish Foundation for the Humanities and Social Sciences. Eriksson also acknowledges funding from the Jan Wallander and Tom Hedelius foundation. We are grateful for valuable comments from participants at various meetings and seminars, including RC28 (New York City), EHES (Tübingen), IPC-IUSSP (Cape Town), EHS (Keele), NBER-Cohort studies (UCLA), Southern Danish University Odense, Umeå University, Stockholm University, Uppsala University, SSB (Oslo), and IFN (Stockholm).

Abstract

Contemporary Western countries show a strong socioeconomic gradient in health and mortality, but it remains unclear if such a gradient existed in historical societies as well. We use linked full-count censuses and death registers for Sweden, covering the cohorts 1841-1920, to study the development of the socioeconomic differences in longevity across these cohorts. Initially men in white-collar occupations had shorter lifespans than working class men and farmers. The modern positive gradient does not become evident until the 1950s. For women, a modern gradient is apparent in all cohorts although differences were much smaller than today.

Keywords: Socioeconomic status, Inequality, Mortality, Longevity

The positive association between socioeconomic status (SES) and life expectancy in modern developed countries is a well-established fact (Mackenbach 2019). Whether measured by income, education or occupational status, SES is positively associated with health and negatively associated with all-cause mortality (e.g. Smith 1999, 2004; Marmot 2004; Elo 2009; Torssander and Erikson 2010; Deaton 2016; Mackenbach 2019). Recent estimates of the association between income and life expectancy puts the issue into perspective: life expectancy at age 40 for U.S. men and women belonging to the top income quartile exceed the bottom quartile by 8.6 and 5.4 years, respectively (Chetty et al. 2016). Comparable inequalities in longevity are also present in more egalitarian settings, such as Sweden (Hederos et al. 2018) and Norway (Kinge et al. 2019). These disparities are the result of almost 50 years of divergence in life expectancies between the top and the bottom, primarily caused by underlying differences in cause-specific mortality in smoking-related cancers and cardiovascular diseases (Hederos et al. 2018; Kinge et al. 2019; Mackenbach 2019).

While the current health inequalities are well documented and understood, the same cannot be said for the period before the 1960s. It is often argued that SES differences in mortality in the past were at least as large as they are today (e.g. Antonovsky 1967; Smith 2004). In a review of the literature, Elo (2009) states that SES is negatively and systematically correlated with mortality in all societies where it has been studied. The reason, according to Link and Phelan (1996), is that high SES groups always have been able to avoid premature deaths regardless of the specific mechanisms and prevailing disease patterns (see also Clouston et al. 2016). The empirical support for the historical validity of these claims is, however, mixed (see Bengtsson and Van Poppel 2011 and Bengtsson, Dribe, and Helgertz 2020 for a summary of the historical evidence). Some studies even suggest that current mortality differentials are of a very recent origin, only emerging in the post-World War II period (Vågerö and Norell 1989; Kunst et al. 1990; Bengtsson, Dribe, and Helgertz 2020; Debiasi and Dribe 2020).

The aim of this paper is to provide population based individual-level estimates of the SES gradient in longevity for much of the twentieth century. By combining micro-level census data and death records, we construct a comprehensive sample of Swedish men and women born 1841-1920, cohorts that to varying degrees experienced and benefited from the introduction of the welfare state and modern medical technology. The censuses of 1880, 1890, 1900,1910, 1930 and 1950 provide individual-level information about SES together with demographic and geographic characteristics. We link individuals appearing in the censuses to death records of all decedents in Sweden between 1880 and 2020. The resulting sample constitutes a unique historical source comprising more than 2.6 million men and women, which we follow from age 40 until death.

The comprehensiveness of our data allows for additional analysis and insights pertaining to selection and the generalizability of the results across varying contexts. These are empirical issues especially relevant for historical populations, which we confront directly. Unlike previous studies of the historical SES gradient in adult mortality, we account for the role played by selection caused by early life mortality, an important issue given the high levels of childhood mortality in the nineteenth century, which also had a socioeconomic gradient already before 1900 (Dribe and Karlsson 2021). By creating multi-generational links, we consider the role played by the SES of the preceding generation on adult mortality. Moreover, we account for unobserved heterogeneity shared between siblings by identifying brothers and sisters and estimating sibling fixed-effects models of adult SES and remaining lifespan at age 40. Finally, by breaking down the analysis to the county level, we show that the main results generalize across various demographic and economic contexts.

Our results show that the association between SES and adult lifespan for 1841-1900 male cohorts was negative and the opposite of the modern gradient: white-collar men had the shortest lifespan of all groups of men and women. Only the last cohorts we study, born 1911-1920, display the modern positive gradient between SES and longevity. The results refute the presumption that there is a universal negative correlation between SES and mortality regardless of time and place. For women we do find a positive relationship between SES and longevity for all cohorts born 1841-1920, similar to the pattern prevailing today. However, the historical inequalities in female adult lifespan were much smaller than they are today. The differences between the longest and shortest living groups rarely exceeded 3 years and in most cases they were much smaller.

We do not aim to assess whether this emerging inequality in longevity reflects a causal effect of SES on mortality, as often assumed in contemporary epidemiological research (see, e.g., Mackenbach 2019 for an in-depth discussion), or a reversed effect from health to SES attainment, which has often been stressed by economists (e.g., Smith 1999, 2004; Deaton 2003; Cutler, Lleras Muney, and Vogl 2012;). Empirical studies using quasi-experimental designs have in some cases identified a causal effect of SES (education or income) on health or mortality (e.g., Lindahl 2005; Lleras-Muney 2005; Lundborg et al. 2016), but in other cases found no such effect (e.g., Cesarini et al. 2016).

Taken together, our results clearly show that the strong positive relationship between SES and longevity is of a recent origin and not a universal historical fact, and that, somewhat counterintuitively, the relationship between SES and longevity emerged and strengthened parallel to the development of modern medicine and the expansion of the welfare state. The latter should come as no surprise as it is well established that there is a pronounced health gradient in all Western societies irrespective of welfare state regime (Mackenbach 2019).

I. SES and mortality in the past

The epidemiological transition is crucial to understand the development of SES differences in health and mortality. According to Omran (1971), mortality has gone through three distinct phases. In the first phase ("the age of pestilence and famine"), mortality was dominated by infectious diseases, and fluctuated widely from year to year due to frequent outbursts of epidemics. During the second phase ("the age of receding pandemics"), the dominance of infectious diseases declined, and epidemic outbreaks became rarer. The third phase ("the age of degenerative and man-made diseases") is characterized by low mortality and by an increasing dominance of non-communicable diseases, such as heart disease and different forms of cancer.

Against the backdrop of the epidemiological transition, Link and Phelan (1995, 1996) identify SES as a *fundamental cause* of mortality, favoring high status groups which have always been able to avoid premature deaths, regardless of the specific mortality risks associated with the epidemiological environment. This implies the existence of a stable SES gradient in mortality over the past 200 years which was not subject to economic and demographic trends. Similarly, Antonovsky (1967) argues for the existence of a constant negative relationship between SES and mortality but qualifies the relationship further by making allowances for differences in the severity of the gradient depending on both the specifics of the disease burden and the relative inequality between high and low status groups.

A recent extension of Link and Phelan's (1995) theory by Clouston et al. (2016) focus on the relationship between SES and cause-specific mortality. The theory retains the notion of SES as a fundamental cause in terms of all-cause mortality, but more precisely consider how cause-specific mortality vary by SES as knowledge about diseases and treatments evolves and disseminates through society. According to Clouston et al. (2016) diseases pass through different stages of prevalence and preventative measures, which are associated with SES differences in cause-specific mortality related to the disease. In the first stage of "natural mortality", diseases are largely non-preventable, resulting in small or even reversed SES differences in cause-specific mortality. In the following stage, new knowledge about prevention and treatments is adopted by higher-status groups, causing widening inequalities in mortality.

With a lag, knowledge is disseminated through society and mortality among the lower-status groups also starts to decline and inequalities begin to narrow. In the final stage, knowledge and treatments are universally accessible and evenly distributed throughout the population, leading to the elimination or severe reduction in the prevalence of the disease and the disappearance of inequalities in mortality between SES groups. No more gains can be made, and some diseases are virtually eliminated (e.g. cholera or tuberculosis). In some cases, however, a small disadvantage for low-status groups remains, due to differences in behavior or a lack of resources necessary for eliminating the disease. The crucial point is that the overall pattern of elimination is repeated disease by disease in a parallel fashion. Therefore, at any given time, because the main causes of mortality continuously shifts from old to new diseases, high-SES groups always have an advantage in terms of lower overall mortality.

The historical evidence on SES differences in adult mortality is fragmentary and mixed. Depending on context and the nature of the analyzed data, SES differences in mortality vary a great deal both in terms of direction and magnitudes. Mortality rates from official statistics for England and Wales document a clear negative gradient between SES and adult mortality from 1890 onwards (Logan 1954; Hollingsworth 1981) albeit declining in magnitude for the period 1920-1950 before widening again (Pamuk 1985). Whether the negative pattern generalize to the period before 1890 is probed by Razzel and Spence (2006) who, based on a range of historical sources, argue that the negative gradient between SES and mortality only emerged at the end of the nineteenth century, and that the relationship might have been insignificant or even positive in some areas in nineteenth century England.

Studies based on micro-level data for geographically limited samples address some of the shortcomings of aggregate statistics but suffer from selection, lack of generalizability and limited sample sizes. An early study by Chapin (1924), on mortality differences in Providence, Rhode Island, in 1865, found that non-taxpayers had higher overall mortality than taxpayers.[1] Similarly, Blum et al. (1990) documents substantial differences in favor of high-status groups in remaining life expectancy in Paris in the 1860s (for men and women at age 40), and in Rouen and Geneva in the seventeenth and eighteenth centuries (for men at age 20). While these results indicate that, historically, SES mattered for health and mortality in urban areas, the generalizability is questioned by Ferrie (2003) who finds small differences in adult mortality by SES, and mixed evidence for the correlation between wealth and mortality, in a rural sample

[1] In addition to overall differences, Chapin (1924) also found that non-taxpayer's mortality to be higher in several important causes of death, such as pulmonary tuberculosis, heart disease and respiratory diseases, while there was only small differences for contagious diseases.

based on the 1850 and 1860 U.S. federal censuses. Two recent studies of educational differences in adult longevity based on the U.S. census of 1940 and linked to different death records show a positive educational gradient for cohorts born in the first decades of the twentieth century (Halpern-Manners et al. 2020; Lleras-Muney, Price, and Yue 2022). For Sweden the available evidence show that a clear positive gradient between SES and adult longevity only emerged sometime after World War II (Bengtsson and Dribe 2011; Edvinsson and Broström 2012; Bengtsson, Dribe, and Helgertz 2020; Debiasi and Dribe 2020).

.

II. Context: Late nineteenth and early twentieth-century Sweden

We focus on men and women born between 1841-1920, cohorts who came of age and lived through a period in which Sweden developed from a poor agricultural society to one of the world's richest and most equal industrialized countries (Schön 2010). Although the full development of the welfare state took place later, initial social reforms, including universal pensions, improved health care, and investments in housing and sanitation, were enacted during the first decades of the twentieth century. These improvements did, however, not prevent large segments of the population from living in poor housing conditions with limited access to health care and education in the nineteenth and early twentieth centuries.

The Swedish mortality transition began in the late eighteenth century with a decline in infant mortality, followed by a decline in child mortality. Around the mid-nineteenth century life expectancy started to increase secularly as mortality declined for working-age adults and the elderly. These gains were followed by a period of stagnation around the turn of the century before a new period of improvement began in the 1930s (Hofsten and Lundström 1976).

The transition is illustrated in Figure 1 which plots period life expectancy at age 40 (e_{40}) from 1751 to 2020. Over the last 200 years, adult life expectancy has increased by around 20 years. Female life expectancy exceeded male life expectancy in all years by about 2-3 years until the mid-twentieth century, when a more pronounced female advantage became apparent. In the 1970s and 1980s the female advantage reached its peak of around 5 years before falling back to under 3 years in 2020 (e_{40} = 44.9 for women and 41.7 for men).

Figure 1

From Figure 1 the decline in volatility is also apparent. The first reduction took place in the first half of the nineteenth century and marks the end of the first phase of the epidemiological transition in Sweden, a development consistent with a decline in epidemics and increasing resistance to short-term economic fluctuations (Bengtsson and Ohlsson 1985). The second phase lasted from the beginning of the nineteenth century until the beginning of the twentieth century. Although mortality was still subject to significant annual variation, it was much more limited than before. Throughout the period, the dominance of infectious diseases declined, and epidemic outbreaks became rarer, ending with the outbreak of the Spanish Flu in 1918. In the final phase of Sweden's epidemiological transition, which began in the early twentieth century and is still ongoing, developments in diagnosis and treatment reduced infectious diseases and all but eliminated annual variation and epidemic outbreaks. In its place heart disease and cancer emerged as the dominant causes of death (Willner 2005a).

Figure 2

Figure 2 displays cause-specific mortality rates in Sweden 1911-1940 for infectious diseases, diseases related to the circulatory system, and cancers. Apart from the sharp increase caused by the Spanish Flu in 1918, the decline in infectious disease mortality was steady. By the mid 1920s it had been overtaken as a leading cause by circulatory diseases, and ten years later infectious disease mortality had fallen below cancers in terms of importance. The 1841-1920 cohorts of men and women which we focus on thus came of age and lived through one of the most dramatic periods in Swedish history. The earlier cohorts had reached middle age by the time Sweden started to industrialize in earnest and were subject to an epidemiological context still dominated by infectious diseases. In contrast, the later cohorts came of age when Sweden had become a leading industrialized nation, and one of the most economically equal countries in the world.

III. Data

A. Sources and data structure

We base our analysis on the enumerated population of Sweden 1880-1950 and those deceased thereafter. Information about occupations and demographic and geographic variables are taken from the 1880, 1890, 1900, 1910, 1930 and 1950 censuses[2]. The censuses are standardized to have the same format as the Integrated Public Use Microdata Series (IPUMS). The Swedish Death Index (2021) is a genealogical resource, which includes the names, sex, and place and date of birth and death for all decedents in Sweden 1830-2020. We focus on the cohorts born between 1841 and 1920 and base our measure of SES on occupational information recorded in

[2] The 1880-1930 full-count censuses have been digitized by the Swedish National Archives and published by SwedPop, a national research infrastructure funded by the Swedish Research Council and partner institutions, www.swedpop.se. The version of the 1930 census used here is a 33% sample of the full population and includes the full population residing in parishes which have been completely digitized. All Swedish counties are represented in the sample apart from two (Kopparberg and Gävleborg). The 1950 full-count census has been digitized by Arkiv Digital and the data have been made available by SwedPop.

the census when aged 30 to 39. Consequently, the occupational information for the 1841-1850 cohorts come from the 1880 census, the 1851-1860 cohorts from the 1890 census, the 1861-1870 cohorts from the 1900 census, and so on. Our main outcome variable is remaining lifespan at age 40.

B. Occupational coding and class scheme

We use the HISCLASS-scheme to measure SES (Van Leeuwen and Maas 2011). HISCLASS is based on HISCO coded occupations (Van Leeuwen, Maas, and Miles 2002) and consists of 12 occupation-based classes which are grouped according to economic sector, whether the occupation is manual or non-manual, its skill level and level of supervision.[3] In the main specification we aggregate HISCLASS into five more general classes: white-collar workers (HISCLASS 1-5), manual skilled workers (HISCLASS 6-8), manual low-skilled workers (HISCLASS 9-10) manual unskilled workers (HISCLASS 11-12) and farmers and fishermen (HISCLASS 8, henceforth referred to as farmers). Because few married women reported an occupation in the census, in the absence of an own recorded occupation we use the husband's occupations to measure SES for wives.[4] Apart from farmers, the scheme is clearly hierarchical in terms of income and status with the white-collar group at the top, followed by the skilled, low skilled and unskilled manual groups in descending order. Because farmers are a very heterogeneous group, spanning from subsistence farmers to large landowners, they are more difficult to position within the hierarchy. Farmers are thus best considered a group unto itself, which is not directly comparable to the white collar and manual groups in terms of resources and status.

C. Probabilistic linking

Personal identification numbers were introduced in Sweden in 1947, and therefore available in the 1950 census and for individuals deceased after 1947 in the death index. This makes the linking of individuals between the 1950 census and the death index straightforward. Because the 1880-1930 censuses precede the introduction of modern identification numbers, we rely on probabilistic linking methods to identify individuals in each source. We use birthplace (parish of birth), sex and birthyear as index variables, meaning that individuals are only considered

---

[3] The coding of occupational text strings into HISCO was done by SwedPop, see www.swedpop.se for full documentation of the coding.

[4] Out of married women aged 30-39, only 0.62% reported an occupation in the 1880 census, 1.24% in 1890, 1.05% in 1900, and 2.39% in 1910.

possible matches if these variables are identical in the death index and the censuses. Names, because of spelling variations, require a more forgiving approach. The similarity of names is therefore measured using the Jaro-Winkler algorithm which assigns a score between 0 (no similarity) and 1 (identical) by comparing characters, character pairs and transpositions in text strings. The algorithm adjusts when strings have the same initial characters and accounts for the fact that irregularities are more common in long strings (see Christen 2006 for a more detailed discussion). To be considered a possible match the Jaro-Winkler similarity score between two names must exceed a threshold of 0.85. This threshold was chosen to minimize the number of false positive links and maximize the linkage rate. To gauge the number of false positive links made at different thresholds, the Jaro-Winkler scores from comparisons of the first recorded first name and surname were plotted against the share of links that could be confirmed based on matching on second and thirds name initials (see Eriksson 2015; Dribe, Eriksson, and Scalone 2019).

To improve linkage rates and minimize false positives links, an iterative approach is used which considers the number of names held by individuals. We begin with all individuals that have at least two recorded first names. In order to be considered a link, the Jaro-Winkler score has to meet the threshold condition for the first two first names and the surname, and constitute a one-to-one relationship between an individual in the death index and the censuses. All links classified as true in the first iteration were then removed from the pool of potential links. The remaining potential links were then used in a second iteration which only rely on the first recorded first name and surname.

The Appendix (Table A1) presents a detailed breakdown of the forward linkage rates between the censuses and the death index by census years and age. In terms of the number of links made, the linking procedure performs comparatively well. The linkage rates for the 1841-1880 cohorts are 70.4% for men and 69.5% for women, rates which are comparable to what is typically achieved when linking between Swedish historical censuses (Wisselgren et al. 2014; Eriksson 2015; Dribe, Eriksson, and Scalone 2019) and exceeds rates achieved for North American and British censuses (Long and Ferrie 2013; Bailey et al. 2020; Abramitzky et al. 2021).

Overall, the linked samples resemble the censuses apart from some minor differences (Appendix, Table A2). The shares of white-collar, skilled and low-skilled men and women in the linked samples are near identical to the distributions observed in the censuses. The share of farmers is higher in the linked samples while the share of men and women belonging to the unskilled group is lower relative to the censuses. Linked individuals are also more likely to be married, have children and not be migrants, although these differences are modest. The average remaining lifespan at 40 for the earlier cohorts, calculated from the linked samples of men and women, is somewhat higher than when calculated using the complete death index or the cohort life expectancy estimates provided by the Human Mortality Database (Appendix, Figure A1). The series do, however, track each other closely in terms of trends and annual variations.

D. Sample selection

We limit our analysis to cohorts born between 1841-1920, observed in the censuses when aged 30-39, and linked to the death index.[5] In total the censuses contain 1,913,255 men and 1,971,124 women belonging to these cohorts, of which we are able to identify 78.9% of men and 77.7% of women in the death index, reducing the sample to 1,510,325 men and 1531,540 women. We exclude men with no recorded occupation in the census, and women with no own or spouse occupation, and all individuals with an occupation which cannot be classified according to the HISCLASS scheme, leaving us with final analytical sample of 1,350,573 men and 1,287,278 women, representing 70.6% and 65.3% of the 1841-1920 cohorts observed in the censuses, respectively. Table 1 displays the descriptive statistics of the analytical sample for men and women separately.

Table 1

IV. Estimation

Our primary outcome of interest is adult longevity, which we measure by the lifespan after age 40. Because the sample only includes extinct cohorts, observations are not subject to censoring. The average number of years lived after age 40 by a cohort is equivalent to cohort life expectance at age 40 (cohort e_{40}). To account for a number of confounders, the relationship between SES and lifespan after age 40 is estimated using the following linear fixed-effects regression model:[6]

$$y_{ij} = \alpha + \beta SES_i + \beta Cohort_i + \beta SES_i Cohort_i + \beta X_i + \partial_j + \epsilon_{ity} \qquad (1)$$

where y_{ij} is the number of years lived after age 40 for individual i in place of residence (county or parish) j, SES_i is the socioeconomic status of the individual, $Cohort_i$ denotes 5-year birth

[5] The sample selection criteria imply that individuals of later cohorts reaching more than 100 years of age are excluded together with permanent emigrants.

[6] See de la Croix and Licandro (2015), and Cummins (2017) for similar approaches.

cohorts, X_i is a vector of individual control variables including marital and migrant status and the number of own children in the household. The analytical sample includes cohorts born between 1841 and 1920 whose lives spanned a period of substantial economic growth, rising wages, continuous urbanization and increasing investments in public health. To allow for changes in the association between SES and mortality over time we include an interaction term between SES and birth cohort.

Regional differences in mortality were often large in the past due to a multitude of factors including population density, communication networks, access to sanitation and safe water, organization of poor relief and health care, breast-feeding practices, and differences in agricultural productivity (Edvinsson and Lindkvist 2011; Garrett et al. 2001; van Poppel, Jonker, and Mandemakers 2005). Because more skilled and well-paid occupations tended to be concentrated in more densely populated areas with higher mortality rates, omitted environmental variables are a likely source of estimated differences in life expectancy by SES. We address this concern by estimating life expectancy net of unobserved local characteristics captured by a series of fixed effects (∂_j) which account for heterogeneity at the county of birth (24 counties) and parish of residence (2542 parishes) levels.

V. Results

A. Descriptive results

There are notable differences in the unadjusted lifespan for the different SES groups (see Table 2). White-collar men belonging to the earliest observed cohort (1841-1845) had the shortest lifespans after 40 of all groups, on average living for 28.9 years, around 1-2 years shorter than the manual groups, and 3.5 years shorter than farmers belonging to the same cohort. For the last observed cohort (1916-1920) the pattern is reversed. The remaining lifespan of the white-collar group now exceeds the manual groups by at least a year, and only trails farmers by 6 months. For women, differences between groups are much smaller, only differing with about a year between groups, with only a weak positive gradient apparent by the difference between the white-collar and manual groups of 0.5-1 years. The differences do, however, become more pronounced for later cohorts and almost double in magnitude by the end of the study period. The basic descriptive pattern thus suggest that the modern positive gradient (albeit weak) was established for women born in the middle of the nineteenth century, but did not appear for men until cohorts born in the beginning of the twentieth century.

Table 2

B. Main results

Figure 3 presents the results of estimations of our main model in the form of predicted lifespan after age 40 for men and women. The basic pattern observed in Table 2 remains after controlling for both demographic and geographic confounders. Overall, absolute differences in lifespan remained stable for men born 1861-1900, with all groups following a common trend of increasing longevity. The stability of the gradient for cohorts born before 1900 stands in contrast to the sudden transition to a modern gradient: for the 1911-1920 male cohorts, the positive association between SES and lifespan is clearly visible.

Figure 3

The estimates for women reveal a different pattern than that observed for men. A clear positive, although not steep, gradient in longevity from the highest status to the lowest is apparent for all cohorts. Women in the white-collar group consistently have the longest lifespan after age 40, while those in the unskilled group have the shortest, with a difference between the two groups of around a year for the oldest cohorts. Beginning with the 1866-1870 cohorts, the groups start to diverge, resulting in an increase from 1 to 2 years difference in lifespan between the top and bottom.

In summary, although all groups benefited throughout the period from increasing longevity, the relatively superior improvement of later cohorts of white-collar men and women provides a precursor to the modern positive relationship between SES and longevity.

C. Selection on family background

How much of the estimated differences in lifespan between SES groups may be attributed to selection? Bias to estimates of SES differences in adult mortality arising from early life conditions is a particular concern since parental resources and child health are closely linked to both adult SES and life expectancy (Case, Lubotsky, and Paxson 2002; Cutler, Lleras-Muney, and Vogl 2012). The expected extent and direction of the bias is, however, not clear *a priori*. Being subjected to adverse conditions in early life may result in scarring that affect later-life outcomes negatively, or selection caused by the survival of healthier individuals (Elo and Preston 1992; Preston, Hill, and Drevenstedt 1998). These mechanisms, depending on the relative strength of each, may thus bias our results in opposite directions.

The positive association between parental SES and child health is evident when examining late nineteenth-century Swedish infant and child mortality: low-status groups had significantly higher infant mortality compared to high-status groups (Edvinsson 2004; Molitoris and Dribe 2016; Dribe and Karlsson 2021;). These differences in childhood conditions imply that bias becomes more severe for lower SES groups, and is positive if selection is the dominant mechanism, and negative if scarring dominates. We address this issue by considering family background and unobserved heterogeneity shared between siblings. Using links between the death index and earlier censuses, we can locate and identify the SES of fathers for 310,632 men and 291,938 women. Similarly, we connect 143,487 men to at least one brother, and 128,717 women to at least one sister within the analytical sample.[7] Because individuals must be observed in their parental home to identify fathers and siblings, our analysis is by necessity limited to considering cohorts which can be observed in a preceding census.

We address selection by first examining the relationship between father SES and lifespan independently, and whether our estimates of the association between own SES and lifespan is sensitive to controlling for father SES. Furthermore, we estimate brother and sister fixed effects models which net out the effect of parental background and exposure to conditions in childhood shared between siblings.

Table 3 and Table 4

Table 4 presents separate regression estimates for 10-year cohorts of men and women including both own and father SES. Estimates of the association between father SES and lifespan mirror estimates for own SES. When controlling for both own SES and father SES, the gradients remain for both variables. Importantly, for men, the shift in the gradient across cohorts from negative to positive and the negative SES gradient for earlier cohorts, is hardly affected by the inclusion of father SES, a finding which speaks against selection being an important explanation for the observed association between adult SES and lifespan or the transition from a negative to positive gradient. For women, results are similar in terms of the impact of controlling for father SES. Across all cohorts and specifications, a distinct advantage for women with both a white-collar father, and white-collar status themselves is evident.

The established gradients between SES and lifespan for both men and women remain after adding sibling fixed effects, although some precision is lost the gradients are still evident, although less steep. Father SES and characteristics shared between siblings thus explain some, but not all the observed relationship between adult SES and lifespan after age 40. That the sibling fixed-effects model yields results that are consistent with our main findings is reassuring, since it implies that for the remaining male gradient to be attributable to negative selection among brothers, frailer men must have been more likely to achieve a higher status

[7] Because the earliest census is 1880 and the identification of fathers and siblings requires that individuals are observed in their parental home in a preceding census these samples only include cohorts born 1861 or later.

than their healthier brothers, or alternatively, healthier men must have been more likely to achieve a lower status than their frailer brothers, prospects which seems highly improbable.

D. Income and life expectancy

In addition to occupational information, the 1930 census also includes individual-level information about taxed income. We use this information to validate our results for SES against observed income in 1930. Moreover, we construct occupational income scores which we apply across all censuses. The scores are constructed by calculating the mean income of all occupations (defined as distinct occupations by their HISCO-codes) for men and women separately.

Table 5

We begin by considering the 1930 census separately to evaluate whether our findings for SES are reflected when estimating the association with lifespan. Because a significant number of individuals reported an income of 0, we assign these individuals to their own category and assign all remaining individuals into quartile groups based on observed incomes. The results are presented in table 5. Although individuals with no reported income have the lowest estimated life expectancy (almost 2 years lower than the reference group for men, and about 1 year lower for women) the negative gradient for men, and the positive gradient for women, is apparent for individuals with a reported income. The results hold up after adding SES as a control variable in the model, and even remain after estimating sibling fixed-effects models.[8]

Figure 4

The results using the occupational income scores across all censuses and cohorts offer a similar picture as that provided by estimates based on SES. Although differences between the income quartile groups are small there is a clear disadvantage in terms of lifespan for high income (the highest quartile) male cohorts born 1841-1900, which eventually turns into an advantage for the 1911-1920 cohort. For women there are no discernible differences between the income groups for cohorts born before 1865. Beginning with the 1866-1870 cohort, a divergence between the income groups is evident, resulting in a difference of 2 years in life expectancy between the highest and lowest income groups for the 1911-1920 cohorts.

E. Geographic heterogeneity

---

[8] Similar findings were made in a study of a regional sample in southern Sweden, which showed that a full income gradient in adult mortality did not appear until after 1950, see Debiasi, Dribe, and Brea-Martinez (2023).

How generalizable are these findings across different geographical contexts? The control strategy pursued so far accounts for environmental confounders through fixed-effects models estimated within narrow geographic contexts. We proceed by examining whether the life expectancies of the five SES groups vary by specific geographic conditions.

We estimate separate models for every Swedish county, yielding 24 county-specific estimates of the relationship between SES and lifespan for men and women by cohort. The regression coefficients for SES (with skilled manual as the reference category) are presented in Figure 5. Overall, for both men and women, the directions of the county-specific gradients conform to the main findings. For men, the white-collar group displays a consistent disadvantage relative to other groups across counties for cohorts born before 1880. A clear modern positive correlation between men's SES and lifespan is not discernible until the 1911-1920 cohorts. For women a small positive gradient is apparent across Sweden for cohorts born before 1870. For later cohorts the gradient becomes steeper, and for the last cohorts observed, the advantage of white-collar women is clear across all of Sweden. Although there is some variation in the strength of the relationship between counties, the magnitudes are relatively modest: the estimated gap in lifespan after age 40 relative to the skilled manual group is typically less than 2 years. The consistency of the gradients across counties is particularly striking given the differences in economic conditions and underlying mortality across Sweden at the end of the nineteenth and the beginning of the twentieth century. These findings support the conclusion that our results generalize across various geographic and economic contexts.

Figure 5 and 6

V. Sensitivity tests

In this section we subject our results to several robustness checks. We begin by considering alternative model specifications and outcomes. Thereafter we examine the importance of the specific class scheme employed and women's status before evaluating whether the quality of the linked sample has an impact on the results.

A. Alternative model specifications and outcomes

To ensure that the results hold up when considering longevity at more advanced ages we also use as the remaining lifespan at ages 50, 60, and 70 as alternative outcome variables (see Appendix, Table B1). The results are also robust to estimating alternative statistical models of the relationship between SES and survival. Table B2 presents estimates from accelerated failure time models assuming residuals follow a log-normal distribution. Table B3 presents estimates of probabilities of survival from age 40 until age 50, 60, and 70 respectively. The estimates from the alternative model specifications are consistent with our main results.

B. Alternative measures of SES

For the sake of brevity and clarity the main analysis uses five different SES groups constructed by aggregating up the original 12 classes defined by the HISCLASS-scheme. Figures B1 and B2 displays estimates using the original 12-group HISCLASS classification. The results are in line with those using the 5-group scheme. Farm workers, both lower-skilled and unskilled, enjoyed an longer lifespan, although not as much as farmers; a result indicating that either living on a farm or working with farming had positive implications for longevity. For women there is a clear gradient in the SES differentials even if it is not perfect. The highest-status group, higher managers, has the longest lifespan and the lowest-status group, unskilled workers, have the shortest. Moreover, within the white-collar group there are differences corresponding to SES. Similarly, among non-farm manual workers lifespan vary by skill-level.

In our main model, women's SES is based on own occupation, or in its absence, the spouse's occupation. Table B4 presents estimates separately for women with SES based on own occupation, and SES based on spouse occupation. They show quite marked differences in the estimates for older cohorts (women born before 1871), but fairly similar estimates for the younger cohorts. For the older cohorts, it is especially women with unskilled occupations that stand out with much shorter lifespans than other women with a recorded occupation (2-3 years), but also the longer lifespans of the white-collar group are only visible form women where SES is measured by the spouse's occupation.

C. False links and measurement bias

Probabilistic matching unavoidably produces samples that includes some share of false positive links, resulting in measurement error and biased estimates. Because the Swedish censuses and the death register are of better quality than the British and North American censuses of the period in terms of share enumerated, accuracy and detail (see Eriksson 2015) we expect the rate of false positive links and associated bias to be less of a problem than what is typical for similar linked samples. Yet, to ensure that our results are not driven by the quality of the linked sample, we extract subsamples of links which fulfill specific criteria chosen to minimize false positive links. We limit the sample to individuals who were linked based on information about two first names and excludes all individuals whose names did not match exactly (i.e. we only retain individuals with a Jaro-Winkler score of 1.0 for all names) between the censuses and the death register. Estimates are presented in Table B5 and show similar results as the main findings.

VI. Discussion

Our demonstrate the existence of a negative male and positive female relationship between SES and lifespan after age 40 at the end of the nineteenth and beginning of the twentieth century in Sweden. For women the pattern resembles the modern SES gradient with white-collar women having the longest lifespan, although differences between groups were relatively small. Among men, farmers experienced the longest lifespan after age 40 initially, followed by manual workers, while white-collar workers had the shortest lifespan. Similar observations have been made in local studies of different Swedish contexts for the same period (Edvinsson 1992; Bengtsson, Dribe, and Helgertz 2020, Debiasi and Dribe 2020). Moreover, a study of white-collar workers at General Electric in Schenectady, New York, observed in the 1930s found that the highest-ranked employees had the shortest lifespans (Nicholas 2023). This raises the question of why the most privileged group with regards to status and resources, were also the most disadvantaged group in terms of survival?

A more detailed analysis of differences in cause-specific mortality is the obvious starting point when seeking to explain the generally observed patterns between SES and overall mortality. Because the Swedish Death Index does not include information about causes of death, it is not possible to estimate SES differences in cause-specific mortality for our linked individual-level sample. In Appendix Table A3, we therefore turn to aggregate statistics on male cause-specific mortality for different occupational groups published by Statistics Sweden (1911, 1921, 1931). To a large extent the occupational groups overlap with the SES groups defined in our main analysis. The farming group closely resembles the HISCLASS group of farmers. The industry and mining, and trade and transport groups primarily consist of manual workers, but also include a small number of managers, owners and administrators which would fall into the white-collar group if classified according to HISCLASS. The public service group is more narrowly defined and closely resembles the white-collar group, consisting of priests, teachers, doctors, and civil servants.

Two causes of death set the public service group apart from the rest: circulatory diseases, and diseases related to the nervous system and sensory organs (which includes stroke). In 1931, 18.9 % of all death among the public service group was attributable to circulatory diseases, exceeding the rates for farmers (12.1%), industry and mining workers (13.1%) and trade and transport workers (13.5%) by a wide margin. In line with the overall trends displayed in Figure

2, the share of deaths caused by circulatory diseases increased for all groups between 1911 and 1931. Although differences narrowed relative to other groups, public service workers remained the group most affected by circulatory diseases.[9] Public service workers were thus forerunners in the sense that cause-specific mortality by the end of the nineteenth century already resembled the modern mortality pattern characteristic of the final stage of the epidemiological transition.

Circulatory diseases are sensitive to lifestyle choices, most notably smoking, diet, exercise, and alcohol consumption. In a contemporary context these factors are often cited as important reasons why low SES is related to poor health and higher mortality (Cavelaars et al. 2000; Elo 2009; Adler and Stewart 2010; Mackenbach 2019). Before high SES became associated with the adoption of healthy habits, there are indications that high status had an opposite effect on behavior, enabling individuals to lead a more sedentary life and consume larger quantities of tobacco, rich foods and alcohol.

Smoking was primarily a habit of the middle and upper classes of white-collar men at the turn of the century in Sweden. It was also a social habit, taking place in the restaurants and bars where higher status men socialized, adding emissions from passive smoking to that of active smoking. Working class men, especially in rural areas, used snuff (*snus*), a moist tobacco product placed between the lip and gum (Nordlund 2005). Even though not harmless, the health effects from consuming snuff are much smaller than from tobacco smoking (e.g., Gartner et al. 2007; Lee 2013). It was not until after World War II, when cigarettes became increasingly popular, that social and gender differences in tobacco consumption converged. When the adverse health effects of smoking became universally appreciated in the 1950s and 1960s, the middle/upper classes were the first to stop, which gave rise to the now familiar negative correlation between smoking and SES, not only in Sweden but in other countries as well (Haenszel, Shimkin, and Miller 1956; Lopez, Collishaw, and Piha 1994; Nordlund 2005).

Because the risks associated with tobacco consumption were still unknown, the public debate and policy during the nineteenth century primarily focused on the effects of alcohol, and the consumption of the lower classes (Willner 2005b). Consequently, the temperance movement became closely tied to the labor movement and the working class, a factor which contributed to a decline in the consumption of alcohol between 1870-1920 (Hurd 1994). The introduction of a strict rationing system in 1919 which based allocations on gender, marital

[9] This result is similar Logan (1954) who shows that the coronary heart disease mortality rate of men belonging to the English professional class (the highest status group) in 1930-32 was decidedly higher than the rates of the other classes.

status and SES served to further limit the alcohol consumption of women and the working class (Bruun and Frånberg 1985).

Although relatively few people were overweight in nineteenth century Sweden, the translation and publication of pamphlets and books on dieting are evident of a growing concern and awareness about obesity. Particular attention was paid to providing advice on weight loss that were complementary to the diet and customs of the bourgeoisie (Bildtgård 2002; Nilsson 2011) since obesity were "especially common among members of society with the tastes and means to enjoy fine cuisine" (Pfannenstill 1901:28).

Household budget surveys carried out in 1913-1914 provides evidence in support of important differences in consumption between different income groups. The surveys show that wealthier households consumed larger quantities and spent more money on alcohol, tobacco, sugar, and fat, both in absolute terms and relative to total expenditures on food and drink, than poorer households (Socialstyrelsen 1921; see Appendix Table A4). Some items stand out: the highest income households consumed at least 2 times as much spirit, 6 times as much beer and 3 times as much tobacco per consumption unit as the lowest income households. According to a Swedish survey in 1963, high income earners aged 50-69 years (cohorts 1896-1913) were more likely to be daily smokers than low-income earners, while these differences were less pronounced among the men aged 18-49 (cohorts 1914-1945) (Appendix Table A5). The survey also suggests that the association between income and smoking among men aged 50-69 was most clear for workers and weaker for business owners (Appendix, table A6). These patterns are consistent with the mortality patterns observed among the public service occupations (see Appendix, Table A3). Moreover, the low levels of cigarette consumption reported for the agricultural, forestry and fishing sector is consistent with the longer life span observed for farmers.

Although not conclusive, the available evidence indicates that an unhealthy lifestyle and associated circulatory diseases may explain why white-collar men – an otherwise privileged group with high living standards – initially had the shortest lifespan after age 40 of all men and women. These factors, which are the primary causal mechanisms linking low status or income and a shorter life expectancy today, implies a change in the behaviors associated with SES during the twentieth century, leading to the eventual reversal of the SES-mortality gradient.

VII. Conclusion

Despite the affluence of modern societies, inequalities in health and life expectancy by SES are widespread across the developed world. Such health inequalities are also apparent in the most equal societies with well-developed welfare states, granting universal provision of basic needs, such as food, safe housing, and health care (Mackenbach 2019). This study deals with SES differences in adult lifespan in a period before the full development of the welfare state, when large segments of society still lived in what today would be considered deep poverty, while the privileged enjoyed considerable affluence. It was a context where one could perhaps expect pronounced SES differences in longevity. This was also the case for children in Sweden, who in this period suffered higher mortality if they belonged to the lower classes than if they were born into the upper or middle classes (Burström and Bernhardt 2001; Molitoris and Dribe 2016; Dribe and Karlsson 2021). For adults, however, our findings suggest a very different picture.

We clearly show that adult lifespan differed between SES groups, but also that the SES differentials were highly gendered. Among cohorts born before 1900, farmers experienced the longest life after 40, about 2 years longer than the white-collar group of upper and upper-middle classes. Somewhat surprisingly, the blue-collar workers had lifespans in-between the farmers and the white-collar group. Hence, in these cohorts of men, the white-collar group was the most disadvantaged in terms of longevity. For women we instead found a pattern resembling the modern SES gradient with the shortest lifespan in the working class and the highest among white-collar groups (at least when measured by their spouse's occupation). For women, however, the SES differences were smaller than for men. For the younger cohorts of women (born after 1870) the emergence of a modern SES gradient in mortality is apparent, with the white-collar group experiencing more rapid improvement in life expectancy relative to the other groups.

The period under study begins after the decline of infectious diseases as major causes of death. Instead, heart disease and cancer, both diseases which are highly dependent on lifestyle, was gaining in prominence. White-collar men stand out as a group which had the means to consume richer and unhealthier food, and more alcohol and tobacco, while not being constrained by mores to the same degree as women. Coupled with a more sedentary lifestyle compared to manual workers and farmers, this most likely lead to higher rates of obesity, poorer physical fitness, and shorter lifespans. This conclusion is supported by the gender differences in the SES pattern. Among women, smoking was much less prevalent, as was alcohol consumption. Hence, for high-status women more resources did not result in the adverse health behavior adopted by their husbands, which is reflected in their relatively longer lifespan compared to other classes.

The longer life span of farmers could possibly be related to a comparably healthy lifestyle in terms of an outdoor working environment, and less exposure to work hazards from emissions or physical danger experienced by blue-collar workers. In addition, differences in cigarette smoking may also have contributed to the survival prospects of farmers.

Within the blue-collar group mortality differences for men were surprisingly small, but if anything indicates that the poorest group of unskilled workers had better survival than the better situated skilled workers. Among women, unskilled workers had the shortest lifespans of all women. This pattern might be related to the skilled workers having more resources than the unskilled, which allowed them to live a less healthy life in terms of smoking and alcohol consumption. A study of a regional population in southern Sweden confirms these cause-specific mortality patterns (Debiasi and Dribe 2020) but more national-level evidence on actual living conditions and cause-of-death specific mortality across SES groups is required to confirm the hypothesis that lifestyle factors were crucial in explaining the inverse social gradient in mortality among adult men in Sweden in the early twentieth century. Regardless of the precise mechanisms, we can be quite confident that economic resources did not help much in promoting survival in this pre-welfare state context. Not until later, with ever increasing standards of living, improved knowledge of health and disease, and the development of welfare society, did the modern social gradient in health and lifespan become an established fact for both men and women.

References

Abramitzky, R., Boustan, L., Eriksson, K., Feigenbaum, J., & Pérez, S. (2021). Automated linking of historical data. *Journal of Economic Literature*, 59(3), 865–918

Adler, N. E., & Stewart, J. (2010). Health disparities across the lifespan: meaning, methods, and mechanisms. *Annals of the New York Academy of Sciences,* 1186, 5–23.

Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *The Milbank Memorial Fund Quarterly,* 45, 31–73.

Bailey, M., Cole, C., Henderson, M., & Massey, C. (2020). How well do automated linking methods perform? Lessons from US historical data. *Journal of Economic Literature*, 58(4), 997–1044.

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

Bengtsson, T., & Ohlsson, R. (1985). The standard of living and mortality response in different ages. *European Journal of Population*, 1, 309–326.

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

Bengtsson, T., Dribe, M., & Helgertz, J. (2020). When did the health gradient emerge? Social class and adult mortality in southern Sweden, 1813-2015. *Demography,* 57(3), 953–977.

Bildtgård, T. (2002). *Hur maten blev en risk: Medicinens bidrag till regleringen av det svenska ätandet* (PhD Dissertation). Uppsala University.

Blum, A., Houdaille, J., & Lamouche, M. (1990). Mortality differentials in France during the late 18th and early 19th centuries. *Population: An English Selection*, 2, 163–185.

Bruun, K., & Frånberg, P. (1985). *Den svenska supen: en historia om brännvin, Bratt och byråkrati.* Stockholm: Prisma.

Burström, B., & Bernhardt, E. (2001). Social differentials in the decline of child mortality in nineteenth century Stockholm. *European Journal of Public Health*, 11, 29–34.

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic status and health in childhood: the origins of the gradient. *American Economic Review*, 92 (5), 1308–1334.

Cavelaars, A. E. J. M., Kunst, A. E., Geurts, J. J. M., Crialesi, R., Grötvedt, L., Helmert, U., … Mackenbach, J. P. (2000). Educational differences in smoking: international comparison. *British Medical Journal*, 320, 1102–1107.

Cesarini, D., Lindqvist, E., Östling, R., & Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on Swedish lottery players. *Quarterly Journal of Economics*, 131(2), 687–738

Chapin, C. V. (1924). Deaths among taxpayers and non-taxpayers, income tax, Providence, 1865. *American Journal of Public Health, 14,* 647–651.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., . . . Cutler, D. (2016). The association between income and life expectancy in the United States, 2001–2014. *JAMA*, 315, 1750–1766.

Christen, P. (2006). A comparison of personal name matching: Techniques and practical issues. *Joint Computer Science Technical Report Series: TR-CS-06-02*, The Australian National University.

Clouston, S. A., Rubin, M. S., Phelan, J. C., & Link, B. G. (2016). A social history of disease: contextualizing the rise and fall of social inequalities in cause-specific mortality. *Demography, 53,* 1631–1656.

Cummins, N. (2017). Lifespans of the European elite, 800-1800. *Journal of Economic History*, 77(2), 406–439.

Cutler, D. M., Lleras-Muney, A., & Vogl, T. (2012). Socioeconomic status and health: Dimensions and mechanisms. In M. Gleid & P. C. Smith (Eds.), *The Oxford Handbook of Health Economics* (pp. 124–163). Oxford, UK: Oxford University Press.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature*, 41, 113–158.

Deaton, A. (2016). On death and money: History, facts, and expectations. *JAMA*, 315, 1703–1705.

Debiasi, E., & Dribe, M. (2020). SES inequalities in cause-specific adult mortality: a study of the long-term trends using longitudinal individual data for Sweden (1813-2014). *European Journal of Epidemiology,* 35, 1043–1056.

Debiasi, E., Dribe, M., & Brea-Martinez, G. (2023). Has it always paid to be rich? Income and cause-specific mortality in southern Sweden 1905-2014. *Population Studies*, accepted for publication.

De la Croix, D., & Licandro, O. (2015). The longevity of famous people from Hammurabi to Einstein. *Journal of Economic Growth*, 20, 263–303.

Dribe, M., & Karlsson, O. (2021). Inequality in early life: Social class differences in childhood mortality in southern Sweden, 1815–1967. *Economic History Review,* 75(2), 475–502.

Dribe, M., Eriksson, B., & Scalone, F. (2019). Migration, marriage and social mobility: Women in Sweden 1880-1900. *Explorations in Economic History* 71, 93–111.

Edvinsson, S. (1992). *Den osunda staden. Sociala skillnader i dödlighet i 1800-talets Sundsvall.* Stockholm: Almqvist & Wiksell International.

Edvinsson, S. (2004). Social differences in infant and child mortality in nineteenth century Sweden. In M. Breschi & L. Pozzi (Eds.), *The Determinants of Infant and Child Mortality in Past European Populations* (pp. 67–87). Udine, Italy: Forum.

Edvinsson, S., & Lindkvist, M. (2011). Wealth and health in 19th Century Sweden. A study of social differences in adult mortality in the Sundsvall region. *Explorations in Economic History*, 48, 376–388.

Edvinsson, S., & Broström, G. (2012). Old age, health, and social inequality: Exploring the social patterns of mortality in 19th century northern Sweden. *Demographic Research, 26(23),* 633–660.

Elo, I. T. (2009). Social class differentials in health and mortality: Patterns and explanations in comparative perspective. *Annual Review of Sociology, 35,* 553–572.

Elo, I. T., & Preston, S. H. (1992). Effects of early-life conditions on adult mortality: A review. *Population Index*, 58, 186–212.

Eriksson, B. (2015). *Dynamic Decades. A Micro Perspective on Late Nineteenth Century Sweden* (PhD Dissertation). Lund University: Lund Studies in Economic History 72.

Ferrie, J. P. (2003). The rich and the dead. Socioeconomic status and mortality in the United States, 1850–1860. In D. L. Costa (Ed.) *Health and Labor Force Participation over the Life Cycle: Evidence from the Past* (pp. 11-50). Chicago: The University of Chicago Press.

Garrett, E., Reid, A., Schürer, K., & Szreter, S. (2001). *Changing Family Size in England and Wales: Place, Class and Demography, 1891–1911*. Cambridge: Cambridge University Press.

Gartner, C. E., Hall, W. D., Vos, T., Bertram, M. Y., Wallace, A. L., & Lim, S. S. (2007). Assessment of Swedish snus for tobacco harm reduction: an epidemiological modelling study. *Lancet*, 369, 2010–2014.

Haenszel, W., Shimkin, M. B., & Miller, H. P. (1956). *Tobacco Smoking Patterns in the United States*. Public Health Monograph No. 45. Washington DC: Public Health Service Publication/US Government Printing Office.

Halpern-Manners, A., Helgertz, J., Warren, J. R., & Roberts, E. (2020). The effects of education on mortality: Evidence from linked U.S. census and administrative mortality data. *Demography*, 57, 1513–1541.

Hederos, K., Jäntti, M., Lindahl, L., & Torssander, J. (2018). Trends in life expectancy by income and the role of specific causes of death. *Economica,* 85, 606–625.

Hofsten, E., & Lundström, H. 1976. *Swedish Population History: Main Trends from 1750 to 1970*. Stockholm: Liber Förlag.

Hollingsworth, J.R. (1981). Inequality in Levels of Health in England and Wales, 1891-1971. *Journal of Health and Social Behavior*, 22, 268–283.

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

Hurd, M. (1994). Liberals, Socialists, and Sobriety: The Rhetoric of Citizenship in Turn-of-the-Century Sweden. *International Labor and Working-Class History,* 45, 44-62.

Kinge J.M., Modalsli J.H., Øverland S…, & Vollset, S. E. (2019). Association of household income with life expectancy and cause-specific mortality in Norway, 2005-2015. *JAMA,* 321, 1916–1925.

Kunst, A.E., Looman, C.W.N., & Mackenbach, J.P. (1990). Socio-economic mortality differences in the Netherlands in 1950–1984: A regional study of cause-specific mortality. *Social Science & Medicine*, 31, 141-152,

Lee, P. N. (2013). Epidemiological evidence relating snus to health – an updated review based on recent publications. *Harm Reduction Journal*, 10, 36.

Lindahl, M. (2005). Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human Resources*, 40, 144–168.

Link, B. G., & Phelan, J. C. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, 35, 80–94.

Link, B. G., & Phelan, J. C. (1996). Understanding sociodemographic differences in health - the role of fundamental social causes. *American Journal of Public Health, 86,* 471–473.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies*, 72, 189–221.

Lleras-Muney, A., Price, J., & Yue, D. (2022). The association between educational attainment and longevity using individual-level data from the 1940 census. *Journal of Health Economics*, 84, 102649.

Logan W.P. (1954). Social class variations in mortality. *British Journal of Preventive & Social Medicine,* 8, 128-37.

Long, J., & Ferrie, J. P. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review*, 103(4), 1109–1137

Lopez, A. D., Collishaw, N. E., & Piha, T. (1994). A descriptive model of the cigarette epidemic in developed countries. *Tobacco Control*, 3, 242–247.

Lundborg, P., Lyttkens, C. H., & Nystedt, P. (2016). The effect of schooling on mortality: New evidence from 50,000 Swedish twins. *Demography*, 53, 1135–1168.

Mackenbach, J. P. (2019). *Health Inequalities. Persistence and Change in European Welfare States*. Oxford: Oxford University Press.

Marmot, M. (2004). *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. London: Bloomsbury Press.

Molitoris, J., & Dribe, M. (2016). Industrialization and inequality revisited: Mortality differentials and vulnerability to economic stress in Stockholm, 1878-1926. *European Review of Economic History*, 20, 176–197.

Nicholas, T. (2023). Status and mortality: Is there a Whitehall effect in the United States? *Economic History Review* online.

Nilsson, F. (2011). *I ett bolster av fett. En kulturhistoria om övervikt, maskulinitet och klass.* Lund: Sekel.

Nordlund, A. (2005). Tobaksrökning och hälsa i Sverige under 1900-talet. In J. Sundin et al. (Eds.), *Svenska folkets hälsa i ett historiskt perspektiv* (pp. 304–361). Stockholm: Statens folkhälsoinstitut.

Omran, A. R. (1971). The epidemiological transition: a theory of the epidemiology of population change. *The Milbank Memorial Fund Quarterly, 49,* 509–538.

Pamuk, E. R. (1985). Social class inequality in mortality from 1921 to 1972 in England and Wales. *Population Studies, 39,* 17–31.

Pfannenstill, S.A. (1901). *Sjuklig fetma (fettsot) och dess behandling.* Stockholm: Bille.

Preston, S. H., Hill, M. E., & Drevenstedt, G. L. (1998). Childhood conditions that predict survival to advanced ages among African-Americans. *Social Science & Medicine*, 47(9), 1231–1246.

Razzel, P., & Spence, C. (2006). The hazards of wealth: Adult mortality in pre-twentieth-century England. *Social History of Medicine,* 19, 381–405.

Schön, L. (2010). *Sweden's Road to Modernity: An Economic History*. Stockholm, Sweden: SNS.

Smith, J. P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives*, 13, 145–166.

Smith, J. P. (2004). Unraveling the SES-health connection. *Population and Development Review,* 30 Suppl., 108–132.

Socialstyrelsen (1921). *Levnadskostnaderna i Sverige 1913–1914. Del 1. Utredningens huvudresultat*. Sveriges officiella statistik, K. Socialstatistik. Stockholm: Statistics Sweden.

Statistics Sweden (1911-40). *Dödsorsaker år 1911-1940.* Sveriges officiella statistik, folkmängden och dess förändringar. Stockholm: Statistics Sweden.

Statistics Sweden (1965). *Rökvanor i Sverige, en postundersökning våren 1963*. Stockholm: Statistics Sweden.

Statistics Sweden (2016). Livslängd och dödlighet i olika sociala grupper. *Demografiska rapporter* 2016:2. Stockholm: Statistics Sweden.

Torssander, J., & Erikson, R. (2010). Stratification and mortality-a comparison of education, class, status, and income. *European Sociological Review*, 26, 465–474.

Van Leeuwen, M. H. D., Maas, I., & Miles, A. (2002). *HISCO: Historical International Standard Classification of Occupations*. Leuven: Leuven University Press.

Van Leeuwen, M. H. D., & Maas, I. (2011). *HISCLASS. A Historical International Social Class Schem*e. Leuven: Leuven University Press.

Van Poppel, F., Jonker, M., & Mandemakers, K. (2005). Differential infant and child mortality in three Dutch regions, 1812–1909. *Economic History Review*, 58, 272–309.

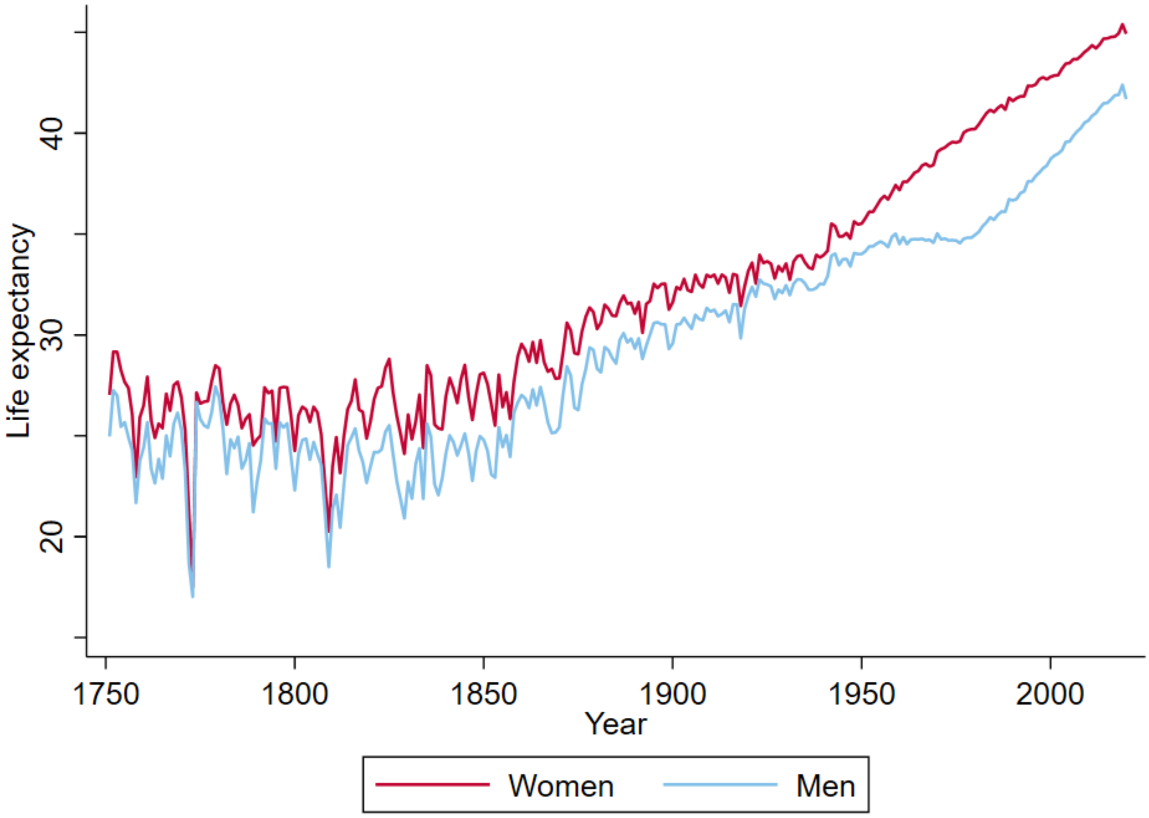
Vågerö, D., & Norell, S. E. (1989). Mortality and social class in Sweden. Exploring a new epidemiological tool. *Scandinavian Journal of Social Medicine*, 17, 49–58.

Willner, S. (2005a). Hälso- och samhällsutveckling i Sverige 1750-2000. In J. Sundin et al. (Eds.) *Svenska folkets hälsa i ett historiskt perspektiv* (pp. 35–79). Stockholm: Statens folkhälsoinstitut.

Willner, S. (2005b). Alkoholpolitik och hälsa hos kvinnor och män. In J. Sundin et al. (Eds.) *Svenska folkets hälsa i ett historiskt perspektiv* (pp. 177–222). Stockholm: Statens folkhälsoinstitut.

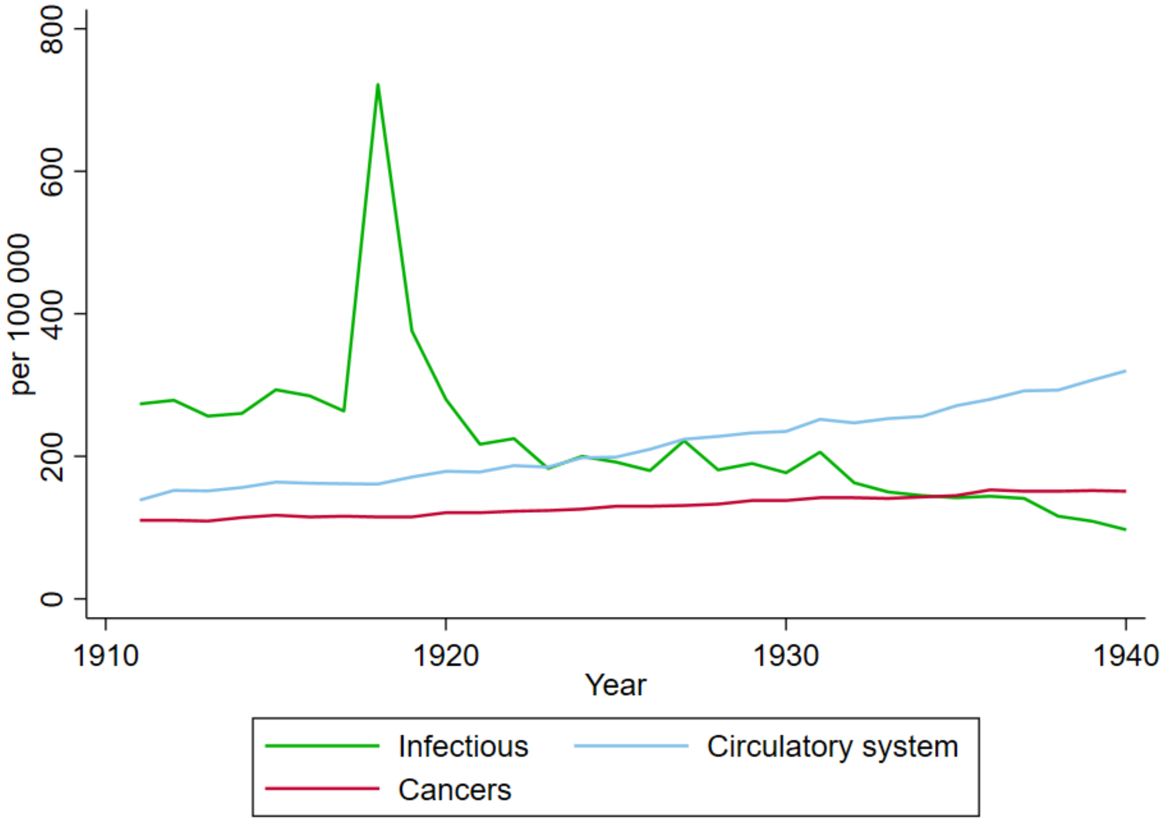
Wisselgren, M. J., Edvinsson, S., Berggren, M., & Larsson, M. (2014). Testing methods of record linkage on Swedish censuses. *Historical Methods*, 47, 138–151.

Figure 1. Period life expectancy at age 40 (e_{40}) in Sweden 1751-2020.



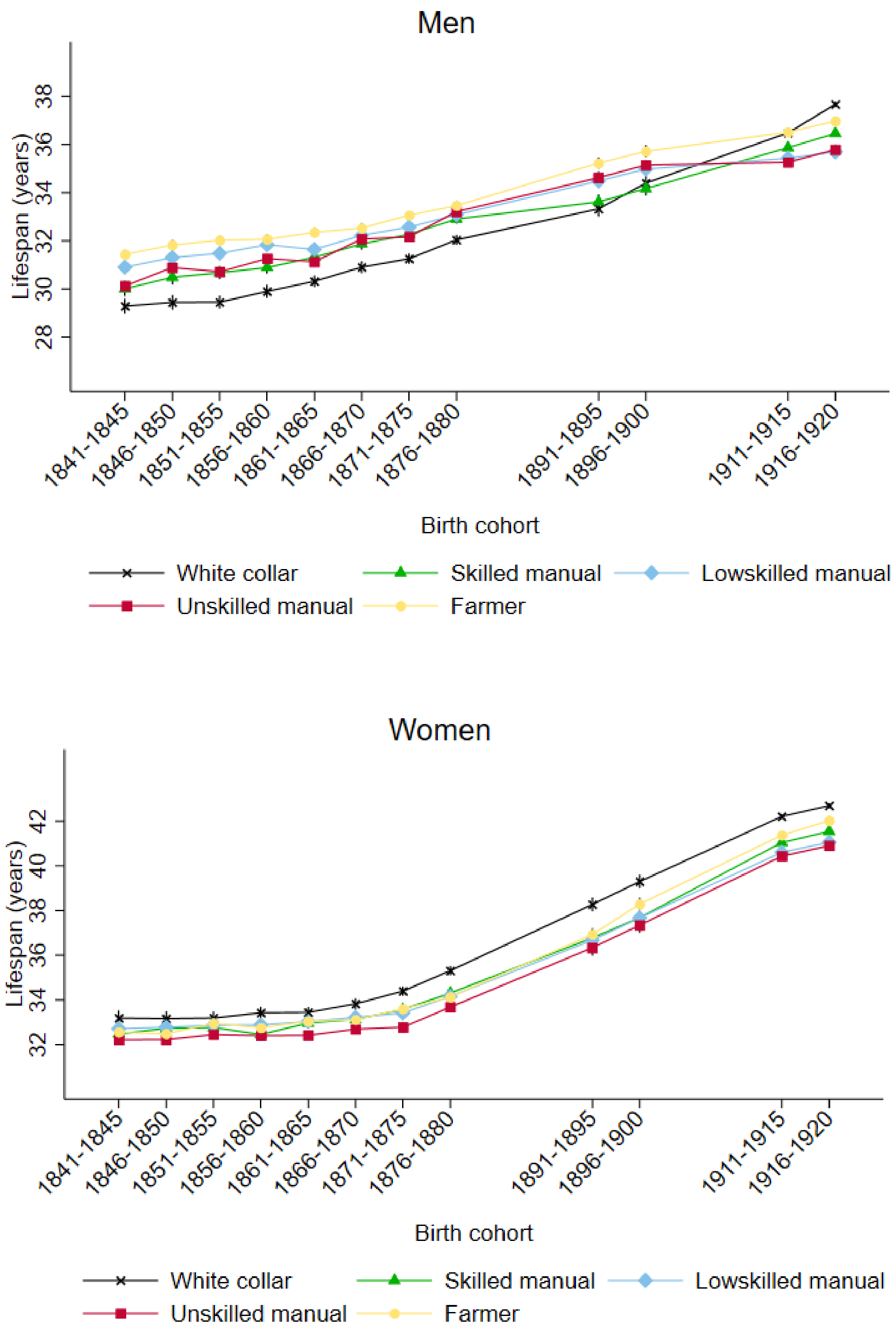
Source: The Human Mortality Database, www.mortality.org.

Figure 2. Cause specific mortality per 100 000, 1911-1940.



Source: Statistics Sweden (1911-1940).

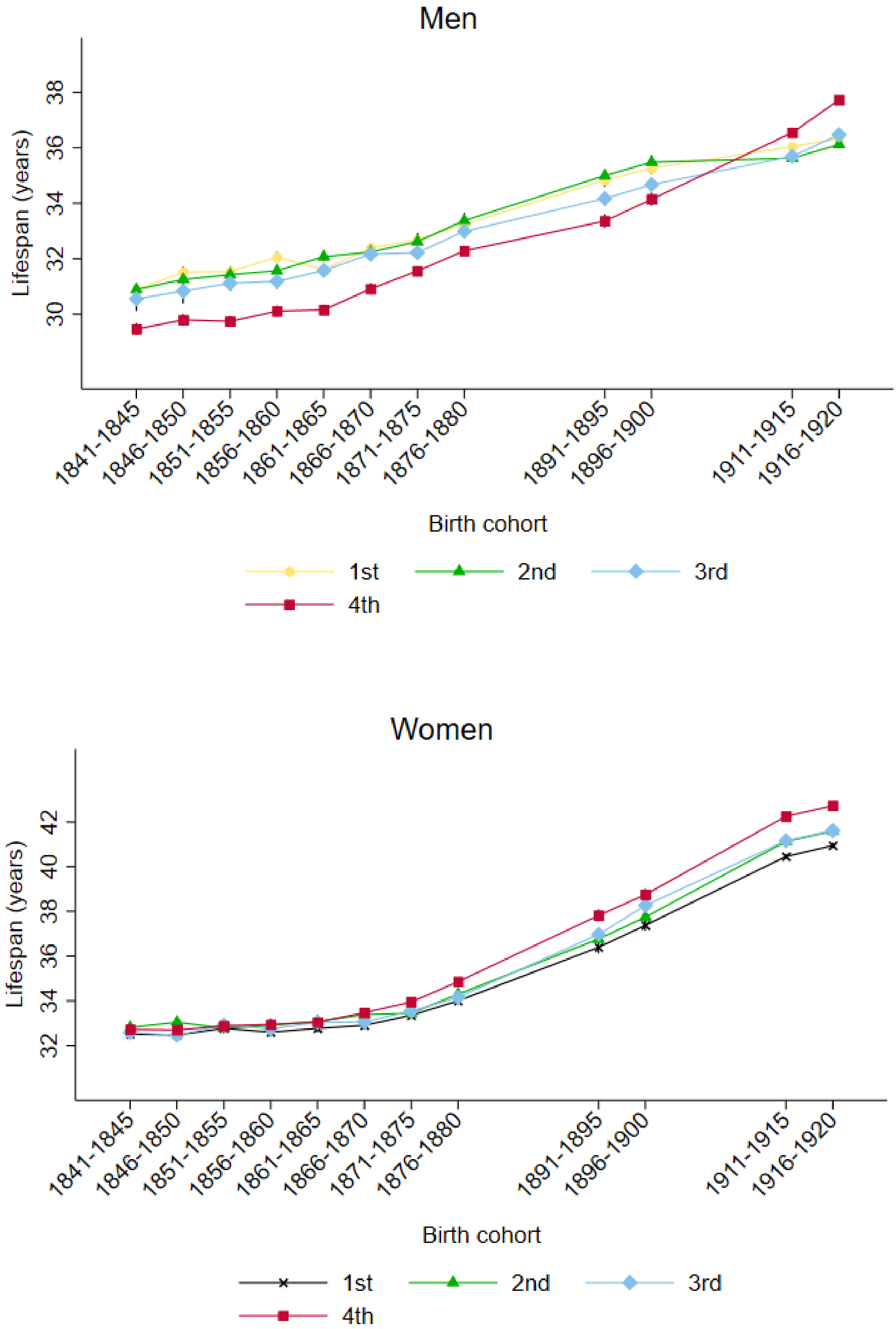
Figure 3. Predicted remaining life span at age 40. OLS estimates, 1841-1920 cohorts (with 95% confidence intervals).



Note: The estimated models control for marital status, children, migrant status and includes fixed effects for county of birth and parish of residence.

Source: See Table 1.

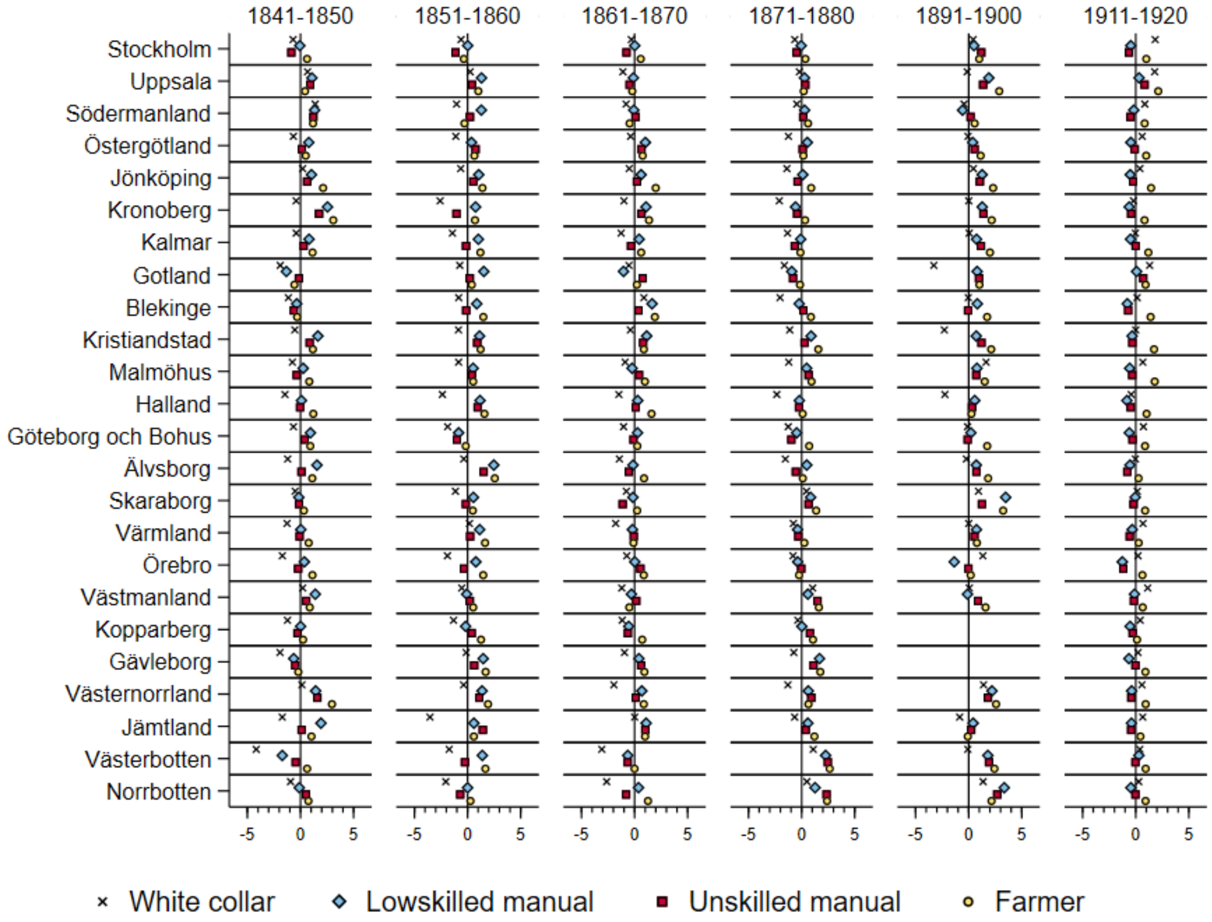
Figure 4. Predicted remaining life span at age 40. OLS estimates, 1841-1920 cohorts (with 95% confidence intervals)



Note: See Figure 3.

Sources: See Table 1.

Figure 5. Coefficients from OLS estimates of remaining life span at age 40 for men, 1861-1920 cohorts.

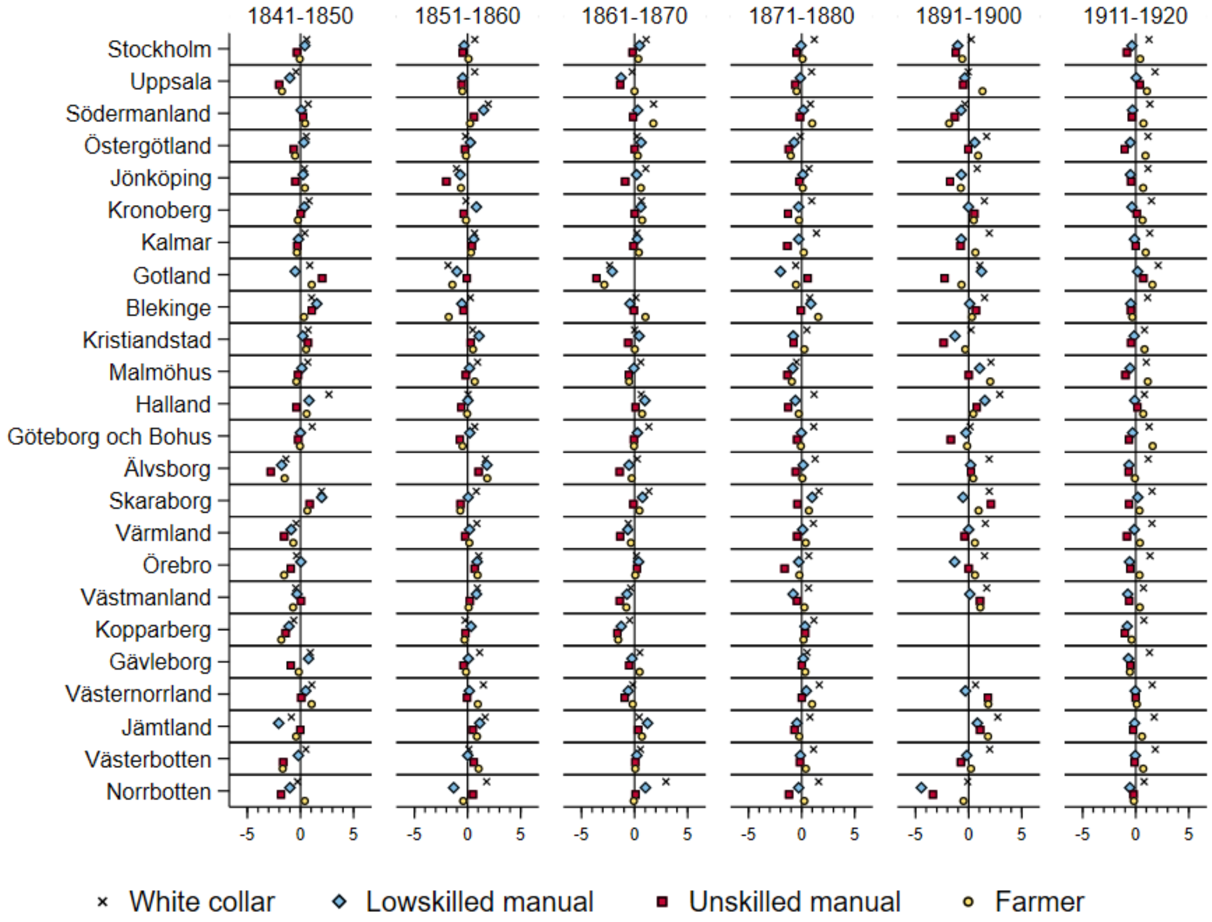


,

Note: See Figure 3.

Sources: See Table 1.

Figure 6. Coefficients from OLS estimates of remaining life span at age 40 for women, 1861-1920 cohorts.



Note: See Figure 3.

Sources: See Table 1.

Table 1. Descriptive statistics of analytical sample.

	1880	1890	1900	1910	1930	1950	All
Panel A. Men							
SES							
White collar	0.11	0.12	0.14	0.15	0.13	0.27	0.18
Skilled manual	0.11	0.13	0.15	0.18	0.15	0.20	0.17
Low-skilled manual	0.20	0.21	0.22	0.24	0.26	0.27	0.24
Unskilled manual	0.30	0.29	0.27	0.24	0.25	0.17	0.23
Farmer	0.29	0.24	0.22	0.19	0.21	0.09	0.17
Marital status							
Unmarried	0.18	0.20	0.22	0.26	0.32	0.22	0.23
Married	0.80	0.78	0.77	0.73	0.66	0.76	0.76
Previously married	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Children							
No children	0.28	0.28	0.30	0.34	0.43	0.38	0.34
One child	0.16	0.15	0.15	0.15	0.20	0.26	0.19
Two children	0.19	0.17	0.17	0.17	0.17	0.23	0.19
Three children	0.16	0.16	0.15	0.14	0.10	0.09	0.12
Four or more children	0.21	0.24	0.24	0.21	0.10	0.04	0.15
Migrant	0.19	0.24	0.25	0.26	0.19	0.31	0.26
Age	34.59	34.48	34.68	34.39	34.36	34.52	34.51
Lifespan after 40	31.14	31.35	31.66	32.52	34.82	36.11	33.63
Observations	146,922	166,918	189,846	237,527	103,363	505,997	1,350,573
Panel B. Women							
SES							
White collar	0.09	0.11	0.14	0.17	0.17	0.27	0.21
Skilled manual	0.10	0.11	0.14	0.16	0.18	0.20	0.15
Low-skilled manual	0.33	0.32	0.31	0.31	0.30	0.27	0.30
Unskilled manual	0.21	0.22	0.21	0.17	0.13	0.17	0.17
Farmer	0.28	0.24	0.21	0.19	0.22	0.09	0.17
Marital status							
Unmarried	0.17	0.18	0.19	0.21	0.23	0.22	0.17
Married	0.82	0.81	0.80	0.78	0.76	0.76	0.81
Previously married	0.01	0.01	0.01	0.01	0.01	0.02	0.01
Children							
No children	0.21	0.22	0.23	0.25	0.29	0.38	0.24
One child	0.15	0.14	0.14	0.14	0.20	0.26	0.20
Two children	0.17	0.16	0.16	0.17	0.20	0.23	0.21
Three children	0.17	0.16	0.15	0.15	0.13	0.09	0.14
Four or more children	0.30	0.32	0.32	0.30	0.18	0.04	0.21
Migrant	0.18	0.23	0.26	0.29	0.23	0.31	0.29
Age	34.54	34.40	34.66	34.40	34.42	34.52	34.49
Lifespan after 40	32.67	32.84	33.05	33.90	37.49	36.11	36.61
Observations	139,785	162,502	185,109	224,250	93,443	505,997	1,287,278

Sources: Census data and the Swedish Death Index (SDI) from SwedPop, www.swedpop.se.

Table 2. Mean remaining life span at age 40 by cohort and SES.

Cohort	White collar	Skilled manual	Low-skilled manual	Unskilled manual	Farmer
Panel A. Men					
1841-1845	28.9	30.0	31.7	30.2	32.4
1846-1850	28.8	30.3	31.9	30.9	32.7
1851-1855	28.9	30.4	32.0	30.6	33.0
1856-1860	29.1	30.5	32.2	31.1	32.9
1861-1865	29.5	31.0	31.9	30.9	33.2
1866-1870	29.8	31.3	32.3	31.8	33.3
1871-1875	30.6	31.9	32.6	31.9	33.9
1876-1880	31.1	32.4	32.9	32.9	34.2
1891-1895	33.2	33.6	34.7	34.6	36.0
1896-1900	34.0	34.0	35.0	35.1	36.3
1911-1915	36.2	35.7	35.4	35.2	37.5
1916-1920	37.2	36.2	35.6	35.8	37.9
Panel B. Women					
1841-1845	33.2	32.6	32.8	32.2	32.7
1846-1850	33.3	32.9	32.8	32.3	32.6
1851-1855	33.2	32.8	32.9	32.5	33
1856-1860	33.5	32.5	32.9	32.5	32.9
1861-1865	33.4	33.0	33.0	32.4	33.1
1866-1870	33.8	33.2	33.2	32.7	33.2
1871-1875	34.4	33.6	33.3	32.7	33.6
1876-1880	35.3	34.3	34.0	33.7	34.2
1891-1895	38.2	36.8	36.6	36.3	37.0
1896-1900	39.2	37.7	37.6	37.3	38.4
1911-1915	42.2	41.1	40.5	40.4	41.5
1916-1920	42.7	41.6	41.0	40.9	42.2

Sources: See Table 1.

Table 3. OLS estimates of remaining life span at age 40 for men, 1861-1920 cohorts.

	1861-1870			1871-1880			1891-1900			1911-1920		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Father SES (ref: Skilled manual)												
White-collar	-0.269	0.0680		-0.325*	0.0275		0.407*	0.501*		0.533***	0.352*	
	(0.202)	(0.213)		(0.146)	(0.152)		(0.199)	(0.203)		(0.156)	(0.159)	
Low-skilled manual	0.930***	0.841***		0.703***	0.612***		0.534***	0.353*		0.212	0.316*	
	(0.161)	(0.165)		(0.117)	(0.118)		(0.149)	(0.150)		(0.128)	(0.129)	
Unskilled manual	0.598***	0.516**		0.449***	0.352**		0.475**	0.302*		-0.238	-0.133	
	(0.165)	(0.169)		(0.116)	(0.117)		(0.152)	(0.154)		(0.145)	(0.146)	
Farmer	1.025***	0.885***		0.769***	0.618***		1.161***	0.832***		1.152***	1.100***	
	(0.156)	(0.165)		(0.114)	(0.118)		(0.145)	(0.151)		(0.123)	(0.127)	
SES (ref: Skilled manual)												
White collar		-0.750***	-0.980*		-0.905***	-0.222		-0.174	-0.514		0.391**	0.0715
		(0.168)	(0.426)		(0.118)	(0.308)		(0.169)	(0.456)		(0.125)	(0.226)
Low-skilled manual		0.303*	0.665		0.230*	0.625*		0.794***	0.723*		-0.501***	-0.449*
		(0.148)	(0.363)		(0.102)	(0.254)		(0.140)	(0.355)		(0.113)	(0.197)
Unskilled manual		0.0931	0.0268		0.168	0.437		0.842***	0.863*		-0.308*	-0.229
		(0.142)	(0.351)		(0.102)	(0.259)		(0.144)	(0.363)		(0.125)	(0.218)
Farmer		0.457**	0.661		0.588***	1.297***		1.319***	1.717***		0.444**	0.335
		(0.157)	(0.388)		(0.116)	(0.289)		(0.155)	(0.390)		(0.146)	(0.251)
R-squared	0.022	0.022	0.133	0.020	0.021	0.127	0.014	0.015	0.099	0.013	0.014	0.102
Observations	100,788	100,788	30,830	174,655	174,655	52,168	86,044	86,044	26,704	105,034	105,034	61,088
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Parish FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Sibling FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Note: Standard errors in parentheses.
Sources: See Table 1.

Table 4. OLS estimates of remaining life span at age 40 for women, 1861-1920 cohorts.

	1861-1870			1871-1880			1891-1900			1911-1920		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Father SES (ref: Skilled manual)												
White-collar	0.747***	0.538*		1.186***	0.937***		0.522*	0.181		0.556***	0.297	
	(0.205)	(0.210)		(0.148)	(0.151)		(0.216)	(0.217)		(0.162)	(0.163)	
Low-skilled manual	0.337*	0.407*		0.152	0.242*		-0.452**	-0.332*		-0.304*	-0.166	
	(0.164)	(0.165)		(0.120)	(0.121)		(0.166)	(0.166)		(0.135)	(0.136)	
Unskilled manual	0.0252	0.134		-0.183	-0.0747		-0.671***	-0.509**		-0.758***	-0.597***	
	(0.168)	(0.169)		(0.120)	(0.120)		(0.169)	(0.169)		(0.154)	(0.155)	
Farmer	0.316*	0.301		0.333**	0.358**		0.131	0.140		0.787***	0.845***	
	(0.159)	(0.162)		(0.119)	(0.120)		(0.162)	(0.164)		(0.129)	(0.130)	
SES (ref: Skilled manual)												
White collar		0.481**	0.433		0.585***	0.285		1.418***	0.714		0.940***	0.572*
		(0.167)	(0.411)		(0.116)	(0.305)		(0.163)	(0.487)		(0.126)	(0.237)
Low-skilled manual		0.0423	-0.283		-0.157	-0.444		0.135	0.189		-0.439***	-0.0250
		(0.145)	(0.353)		(0.102)	(0.261)		(0.142)	(0.375)		(0.121)	(0.219)
Unskilled manual		-0.545***	-0.608		-0.594***	-0.757*		-0.288	-0.325		-0.462**	0.153
		(0.153)	(0.372)		(0.114)	(0.295)		(0.174)	(0.481)		(0.149)	(0.267)
Farmer		0.0775	0.124		-0.156	-0.195		0.462**	0.0540		0.314*	0.376
		(0.158)	(0.389)		(0.120)	(0.304)		(0.163)	(0.437)		(0.154)	(0.271)
R-squared	0.005	0.005	0.124	0.006	0.007	0.098	0.009	0.011	0.089	0.011	0.013	0.082
Observations	98,072	98,072	28,710	168,214	168,214	45,823	76,927	76,927	18,656	91,718	91,718	46,757
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Parish FE	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Sibling FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Note: Standard errors in parentheses.
Sources: See Table 1.

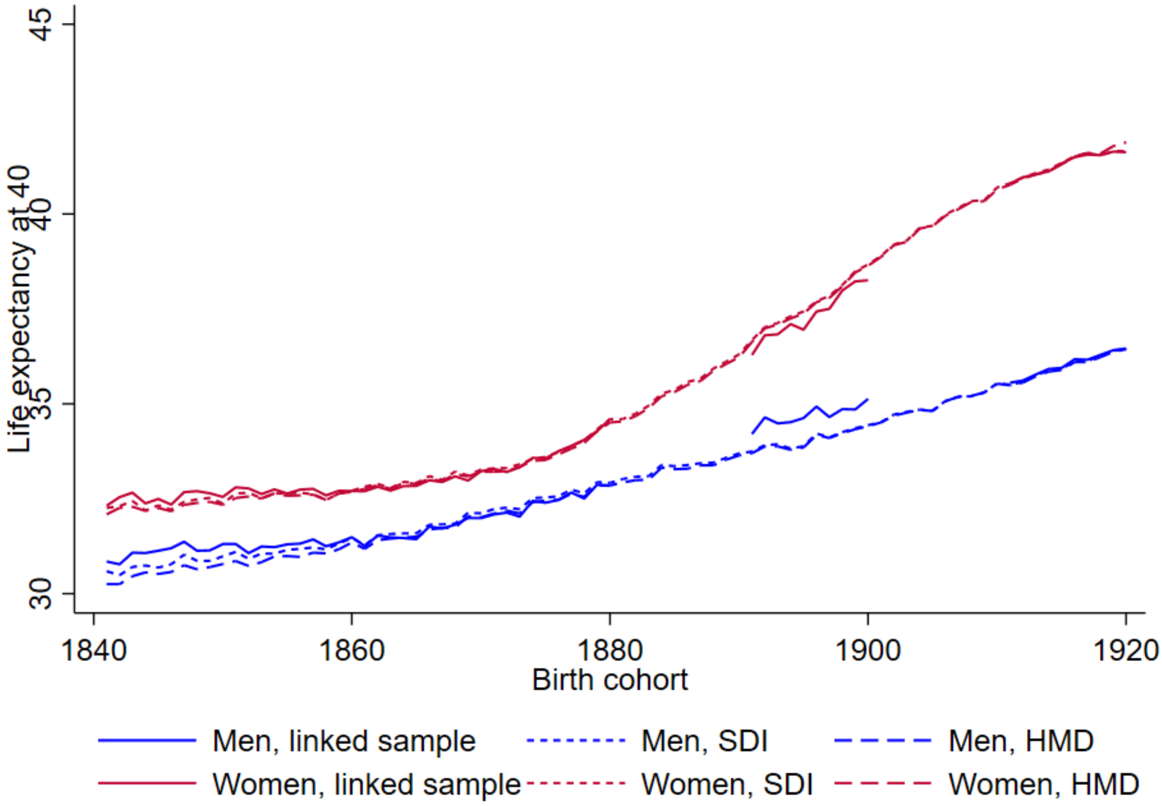
Table 5. OLS estimates of remaining life span at age 40 for 1891-1900 cohorts, 1930 census.

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Income (ref: 2nd quartile group)								
No income/missing	-1.835***	-1.845***	-1.924***	-1.951***	-1.167***	-1.057***	-0.339	-0.292
	(0.153)	(0.153)	(0.410)	(0.408)	(0.160)	(0.160)	(0.460)	(0.462)
1st quartile group	0.190	0.104	0.0410	-0.0254	-0.329*	-0.284*	-0.457	-0.438
	(0.115)	(0.115)	(0.300)	(0.300)	(0.130)	(0.130)	(0.376)	(0.377)
3rd quartile group	-0.185	-0.0982	-0.266	-0.114	0.171	0.0585	0.367	0.335
	(0.116)	(0.117)	(0.313)	(0.315)	(0.128)	(0.128)	(0.376)	(0.378)
4th quartile group	-0.659***	-0.319*	-0.816*	-0.434	0.916***	0.577***	0.313	0.226
	(0.133)	(0.137)	(0.373)	(0.380)	(0.140)	(0.147)	(0.438)	(0.449)
SES (ref: Skilled manual)								
White collar		-0.0882		-0.546		1.160***		0.583
		(0.149)		(0.456)		(0.149)		(0.495)
Low-skilled manual		0.745***		0.616		0.00772		0.205
		(0.126)		(0.355)		(0.129)		(0.377)
Unskilled manual		0.818***		0.802*		-0.369*		-0.317
		(0.130)		(0.365)		(0.157)		(0.484)
Farmer		1.548***		1.686***		0.581***		0.101
		(0.138)		(0.394)		(0.147)		(0.441)
R-squared	0.015	0.017	0.100	0.101	0.010	0.011	0.089	0.089
Observations	103,362	103,362	26,704	26,704	93,443	93,443	18,656	18,656
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	No	No	Yes	Yes	No	No
Parish FE	Yes	Yes	No	No	Yes	Yes	No	No
Sibling FE	No	No	Yes	Yes	No	No	Yes	Yes

Note: Standard errors in parentheses.
Sources: See Table 1.

Online Appendix A: Additional Results

Figure A1. Different estimates of cohort life expectancy at age 40, 1841-1920 cohorts



Notes: SDI: The Swedish death Index: HMD: The Human Mortality database. The estimates for the linked sample and SDI are the average remaining lifespans at age 40, while the estimates from the HMD are based on cohort life tables.

Sources: See Table 1 and the Human Mortality Database, www.mortality.org.

Table A1. Linkage rates by census year and age.

Age	1880	1890	1900	1910	1930	1950	Total
Panel A. Men							
30	60.6	65.2	70.6	74.7	86.5	94.5	77.3
31	61.5	65.3	71.4	74.1	86.5	95.2	76.7
32	62.1	66.8	72.2	75.2	86.6	95.5	77.8
33	63.5	67.3	72.7	75	86.7	96	78.6
34	63.7	68.1	72.6	75.7	86.1	96.1	78.9
35	64	69.1	72.5	75.6	87.1	96.2	79.1
36	64.6	69.3	73.6	76.1	86.3	96.4	79.4
37	65.8	70	73.5	76.4	86.2	96.3	80.1
38	66.3	71.4	74.4	76.8	85.7	96.4	80.8
39	66.6	71.6	75.6	77	85.1	96.6	81.2
Total	63.8	68.2	72.9	75.6	86.3	95.9	78.9
Panel B. Women							
30	59.2	64.8	72.4	76.2	84.6	93.7	76.8
31	58.7	64.8	72.6	76.7	85.4	94.6	76.1
32	59.9	64.9	73.2	76.4	84.6	95.5	76.9
33	60.7	65.9	73.1	76.5	85.9	95.9	77.7
34	61.4	66.6	72.9	76.4	84.9	96.2	78
35	61	66.7	72.6	76.6	85.8	96.3	77.9
36	60.4	66.1	72.8	76.5	85.1	96.7	77.8
37	61.7	67.6	72.6	76.3	85.5	96.6	78.5
38	60.9	67.9	72.6	76.3	85.1	96.8	78.6
39	62.7	67.6	73.1	76.9	85.2	96.8	79.1
Total	60.6	66.2	72.8	76.5	85.2	95.9	77.7

Sources: See Table 1

Table A2. Descriptive statistics for census and linked sample by census year.

	1880		1890		1900		1910		1930		1950	
	Census	Linked	Census	Linked	Census	Linked	Census	Linked	Census	Linked	Census	Linked
Panel A. Men												
SES												
White collar	0.09	0.09	0.1	0.11	0.12	0.12	0.14	0.14	0.13	0.12	0.26	0.26
Skilled manual	0.09	0.1	0.11	0.11	0.13	0.13	0.16	0.16	0.14	0.14	0.2	0.2
Low-skilled manual	0.16	0.17	0.17	0.18	0.18	0.19	0.22	0.22	0.23	0.24	0.26	0.26
Unskilled manual	0.28	0.26	0.27	0.26	0.26	0.24	0.23	0.22	0.23	0.22	0.17	0.17
Farmer	0.22	0.24	0.19	0.21	0.17	0.19	0.16	0.17	0.18	0.19	0.08	0.09
Missing	0.16	0.14	0.15	0.13	0.14	0.12	0.08	0.07	0.09	0.09	0.03	0.03
Marital status												
Unmarried	0.29	0.24	0.31	0.26	0.32	0.28	0.33	0.3	0.36	0.34	0.23	0.23
Married	0.7	0.74	0.68	0.72	0.66	0.71	0.65	0.69	0.62	0.65	0.75	0.75
Previously married	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Children												
No children	0.37	0.33	0.39	0.34	0.41	0.36	0.41	0.37	0.47	0.45	0.4	0.39
One child	0.14	0.15	0.13	0.14	0.13	0.14	0.13	0.14	0.19	0.19	0.25	0.26
Two children	0.16	0.18	0.15	0.16	0.14	0.15	0.15	0.16	0.16	0.16	0.22	0.23
Three children	0.14	0.15	0.14	0.14	0.13	0.14	0.12	0.13	0.09	0.1	0.08	0.09
Four or more children	0.18	0.19	0.2	0.21	0.2	0.21	0.19	0.2	0.09	0.1	0.04	0.04
Migrant	0.21	0.18	0.24	0.22	0.25	0.23	0.27	0.25	0.2	0.19	0.33	0.31
Age	34.35	34.44	34.25	34.34	34.49	34.54	34.26	34.29	34.32	34.31	34.48	34.5
Observations	276,967	176,609	291,291	198,706	306,476	223,381	350,438	264,935	137,712	118,846	550,371	527,735
Panel B. Women												
SES												
White collar	0.07	0.07	0.09	0.09	0.11	0.11	0.14	0.14	0.15	0.15	0.31	0.31
Skilled manual	0.08	0.08	0.09	0.09	0.11	0.11	0.13	0.14	0.16	0.16	0.16	0.16
Low-skilled manual	0.26	0.26	0.25	0.26	0.24	0.25	0.26	0.26	0.25	0.25	0.27	0.27
Unskilled manual	0.17	0.17	0.17	0.18	0.16	0.17	0.14	0.15	0.11	0.11	0.11	0.12

Farmer	0.21	0.22	0.18	0.19	0.16	0.17	0.15	0.16	0.17	0.18	0.08	0.09
Missing	0.22	0.2	0.23	0.2	0.21	0.19	0.17	0.15	0.15	0.14	0.06	0.05
Marital status												
Unmarried	0.29	0.26	0.3	0.28	0.31	0.29	0.32	0.3	0.3	0.29	0.15	0.15
Married	0.68	0.71	0.67	0.7	0.66	0.69	0.66	0.68	0.68	0.69	0.81	0.82
Previously married	0.03	0.03	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.03
Children												
No children	0.3	0.28	0.31	0.29	0.33	0.31	0.33	0.32	0.35	0.35	0.28	0.27
One child	0.15	0.15	0.14	0.14	0.14	0.13	0.14	0.14	0.19	0.19	0.27	0.28
Two children	0.16	0.16	0.15	0.15	0.14	0.15	0.15	0.15	0.18	0.18	0.26	0.27
Three children	0.15	0.15	0.14	0.14	0.13	0.13	0.13	0.13	0.12	0.12	0.12	0.12
Four or more children	0.25	0.26	0.26	0.27	0.26	0.27	0.25	0.26	0.16	0.16	0.07	0.07
Migrant	0.2	0.17	0.23	0.21	0.25	0.24	0.28	0.26	0.24	0.23	0.37	0.36
Age	34.37	34.42	34.24	34.29	34.54	34.54	34.29	34.29	34.35	34.36	34.48	34.51
Observations	298,483	180,850	317,439	210,049	324,920	236,470	357,355	273,301	134,970	114,999	537,957	515,812

Sources: See Table 1.

Table A3. Proportionate cause-specific mortality by occupational groups, men age 15 and above, 1911-1931.

	1911				1921				1931			
	Farming	Industry and mining	Trade and transport	Public service	Farming	Industry and mining	Trade and transport	Public service	Farming	Industry and mining	Trade and transport	Public service
Diseases related to:												
Old age	12.4	6.4	1.3	2.2	13.4	7.2	1.8	2.6	9.3	1.8	0.5	1.0
Infections	18.4	24.7	24.0	19.6	17.2	24.6	23.6	22.1	17.0	21.7	23.9	18.0
Blood system	0.5	0.5	0.7	1.3	0.6	0.8	1.2	1.4	1.0	1.0	1.1	0.8
Chronic poisoning (incl. alcohol)	0.1	0.7	1.0	0.4	0.0	0.1	0.4	0.0	0.0	0.1	0.2	0.1
Metabolic disorders	1.0	1.1	1.8	2.4	1.3	1.1	1.9	1.9	0.9	1.1	1.7	2.5
Nervous system and sensory organs	7.8	6.6	6.5	10.3	8.1	7.4	7.1	8.0	8.8	8.0	7.4	10.0
Mental illness	0.4	0.2	0.6	1.1	0.1	0.4	1.2	1.0	0.3	0.4	0.6	1.2
Circulatory system	11.2	13.1	13.5	18.9	14.8	16.5	16.4	20.6	19.6	20.0	19.4	22.1
Respiratory system	12.7	11.8	8.2	9.4	9.9	9.1	8.1	7.6	8.9	8.7	5.9	7.4
Digestive system	5.0	4.4	5.1	6.2	4.5	4.4	5.2	4.5	4.7	5.3	6.6	7.0
Urinary system	4.9	5.0	5.5	5.4	4.8	5.1	6.1	5.6	5.2	5.5	5.6	7.3
Bone	0.9	0.4	0.4	0.1	0.5	0.5	0.3	0.3	0.2	0.3	0.1	0.0
Skin	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
Tumours	12.0	10.9	10.3	10.2	13.6	12.9	9.9	12.3	14.8	15.1	12.9	13.6
Violence (incl. suicide)	5.6	11.3	19.1	11.1	6.8	7.9	15.5	10.9	7.8	10.0	13.6	9.1
Other and unknown	7.0	2.8	2.2	1.3	4.3	1.8	1.4	1.3	1.6	1.0	0.6	0.0
Total	100	100	100	100	100	100	100	100	100	100	100	100
Number of deaths	8974	5571	2104	756	8261	6931	2205	1012	8317	6217	2661	729

Sources: Statistics Sweden (1915,1926,1935)

Table A4. Yearly consumption of alcohol, tobacco and selected foodstuffs in Sweden, 1913-1914.

	Income per consumption unit (SEK)			
	<600	600-750	750-1050	>1050
	Panel A. Value (SEK)			
Spirits	3.0	4.0	6.4	8.7
Wine	-	-	-	-
Beer	1.0	1.8	3.2	6.0
Tobacco	2.8	4.2	6.1	9.2
Cream	3.6	4.5	5.6	6.4
Wholemilk	26.8	35.9	40.6	49.9
Skimmed milk	7.2	5.6	4.8	2.0
Butter	20.7	26.1	31.3	36.7
Margarine	12.6	11.1	10.4	8.6
Sugar and syrup	19.7	22.2	23.0	25.1
Fat and lard	2.4	2.1	1.9	1.2
	Panel B. Share of total food, drink and tobacco expenditure (%)			
Spirits	1.1	1.2	1.7	2.1
Wine	-	-	-	-
Beer	0.3	0.6	0.9	1.4
Tobacco	1.0	1.3	1.6	2.2
Cream	1.3	1.4	1.5	1.5
Wholemilk	9.6	10.9	11.0	12.0
Skimmed milk	2.5	1.7	1.8	0.5
Butter	7.4	8.0	8.5	8.8
Margarine	4.5	3.4	2.8	2.0
Sugar and syrup	7.0	6.8	6.2	6.0
Fat and lard	-	-	-	-
	Panel C. Volume (litres or kilograms)			
Spirits	2.0	2.5	3.9	4.4
Wine	0.0	0.0	0.1	0.2
Beer	2.6	5.0	8.7	15.8
Tobacco	-	-	-	-
Cream	-	-	-	-
Wholemilk	194.5	257.5	281.1	341.4
Skimmed milk	111.4	86.4	72.1	28.2
Butter	9.6	12.0	14.2	16.7
Margarine	8.8	7.8	7.3	6.0
Sugar and syrup	29.3	33.0	34.7	38.2
Fat and lard	-	-	-	-

Source: Socialstyrelsen 1921.

Table A5. Smoking habits of men in Sweden in 1963.

	Income SEK						
	<5,000	5,000-10,000	10,000-15,000	15,000-20,000	20,000-25,000	25,000-30,000	>30,000
	Ages 18-49 (1914-1945 cohorts)						
Daily smoker	43.5	50.3	54.6	55.8	55.4	56.5	53.8
Occasional smoker	12.5	10.9	9.1	8.0	9.1	9.0	9.1
Non-smoker	44.0	38.7	36.3	36.2	35.5	34.5	37.2
Observations	1200	1226	2188	2414	1432	825	993
	Ages 50-69 (1896-1913 cohorts)						
Daily smoker	31.8	37.8	46.8	53.5	54.2	58.5	50.2
Occasional smoker	16.9	18.4	13.7	9.8	6.8	6.6	9.5
Non-smoker	51.2	43.7	39.4	36.6	38.9	34.8	40.4
Observations	1380	2244	3548	2872	1543	804	1296

Source: Statistics Sweden (1965, Tables 003-004).

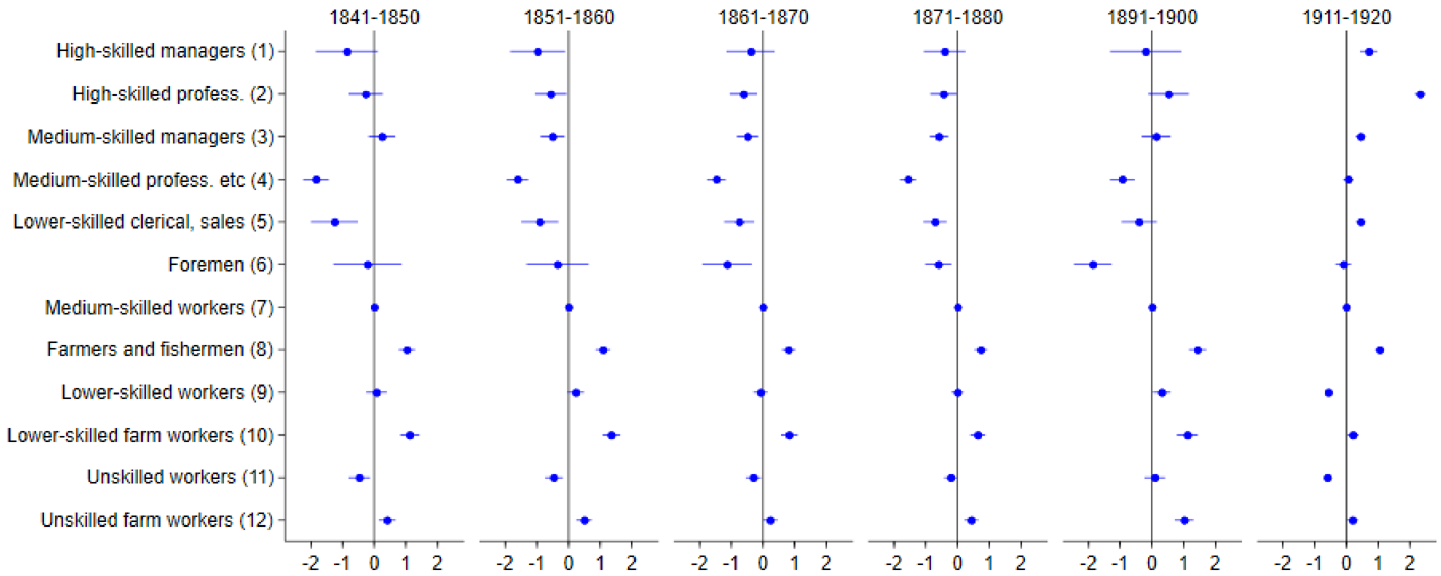
Table A6. Share of men (%) reporting daily consumption of cigarettes in Sweden in 1963.

	Income SEK							
	<5,000	5,000-10,000	10,000-15,000	15,000-20,000	20,000-25,000	25,000-30,000	>30,000	Observations
	Ages 18-49 (1914-1945 cohorts)							
Business owners	43.5	41.3	41.2	41.8	40.0	43.2	40.2	3622
Workers	47.4	48.9	41.5	38.8	42.3	45.9	49.9	4625
Agriculture, forestry and fishing	10.1	8.5	14.8	16.2	15.7	14.3	33.4	1040
	Ages 50-69 (1896-1913 cohorts)							
Business owners	27.3	26.9	29.0	34.3	37.5	39.1	33.6	4076
Workers	18.4	28.3	26.7	31.0	36.7	42.9	44.7	5255
Agriculture, forestry and fishing	3.0	2.3	2.7	3.7	6.4	7.7	18.2	2421

Source: Statistics Sweden (1965, Tables 115-120).

Online Appendix B: Sensitivity testing

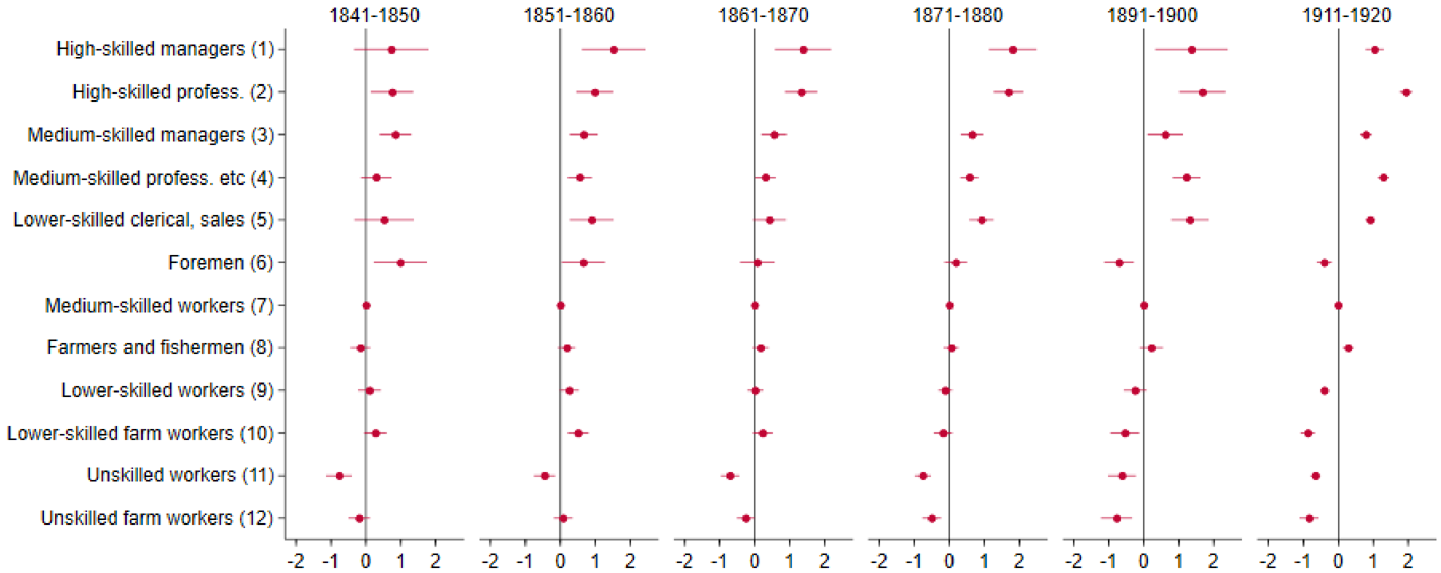
Figure B1. Coefficients from OLS estimates of remaining life span at age 40 (with 95% confidence intervals).



Note: The estimated models control for marital status, children, migrant status and includes fixed effects for county of birth and parish of residence.

Sources: See Table 1

Figure B2. Coefficients from OLS estimates of remaining life span at age 40 (with 95% confidence intervals)



Note: The estimated models control for marital status, children, migrant status and includes fixed effects for county of birth and parish of residence.

Sources: See Table 1

Table B1. OLS estimates of remaining life span at ages 50, 60 and 70, 1861-1920 cohorts.

	1841-1850 (1)	1851-1860 (2)	1861-1870 (3)	1871-1870 (4)	1891-1900 (5)	1911-1920 (6)
Panel A. Men						
Lifespan after 50						
SES (ref: Skilled manual)						
White collar	-0.728***	-1.073***	-0.914***	-0.799***	-0.023	0.835***
	(0.129)	(0.113)	(0.098)	(0.084)	(0.131)	(0.048)
Low-skilled manual	0.452***	0.452***	0.243**	0.338***	0.713***	-0.411***
	(0.114)	(0.100)	(0.088)	(0.074)	(0.111)	(0.047)
Unskilled manual	0.094	0.085	0.094	0.248***	0.691***	-0.342***
	(0.108)	(0.094)	(0.084)	(0.074)	(0.113)	(0.052)
Farmer	0.589***	0.683***	0.574***	0.607***	1.401***	0.875***
	(0.112)	(0.102)	(0.094)	(0.083)	(0.120)	(0.067)
R-squared	0.025	0.025	0.021	0.018	0.014	0.013
Observations	134,301	152,952	174,602	220,395	98,725	490,362
Lifespan after 60						
SES (ref: Skilled manual)						
White collar	-0.523***	-0.782***	-0.663***	-0.447***	0.208	0.853***
	(0.111)	(0.096)	(0.085)	(0.072)	(0.115)	(0.042)
Low-skilled manual	0.282**	0.316***	0.151*	0.226***	0.497***	-0.339***
	(0.097)	(0.084)	(0.075)	(0.063)	(0.097)	(0.042)
Unskilled manual	0.122	0.116	0.190**	0.171**	0.529***	-0.305***
	(0.092)	(0.080)	(0.072)	(0.063)	(0.099)	(0.046)
Farmer	0.392***	0.397***	0.348***	0.444***	1.092***	0.626***
	(0.096)	(0.087)	(0.080)	(0.072)	(0.105)	(0.060)
R-squared	0.019	0.017	0.018	0.014	0.013	0.012
Observations	116,377	132,644	152,872	195,177	90,717	449,919
Lifespan after 70						
SES (ref: Skilled manual)						
White collar	-0.256**	-0.375***	-0.104	0.039	0.429***	0.744***
	(0.095)	(0.083)	(0.074)	(0.062)	(0.100)	(0.036)
Low-skilled manual	0.148	0.229**	0.044	0.104	0.271**	-0.232***
	(0.083)	(0.072)	(0.064)	(0.054)	(0.084)	(0.035)
Unskilled manual	0.040	0.040	0.036	-0.003	0.212*	-0.264***
	(0.078)	(0.068)	(0.062)	(0.054)	(0.086)	(0.039)
Farmer	0.186*	0.203**	0.207**	0.256***	0.670***	0.344***
	(0.082)	(0.074)	(0.069)	(0.061)	(0.091)	(0.051)
R-squared	0.013	0.012	0.012	0.009	0.008	0.011
Observations	86,752	100,059	115,204	149,627	72,648	360,720
Panel B. Women						
Lifespan after 50						
SES (ref: Skilled manual)						
White collar	0.150	0.291*	0.309**	0.607***	1.334***	1.140***
	(0.138)	(0.116)	(0.101)	(0.085)	(0.128)	(0.048)
Low-skilled manual	-0.064	0.192	0.129	-0.109	0.006	-0.357***
	(0.116)	(0.100)	(0.088)	(0.075)	(0.113)	(0.049)
Unskilled manual	-0.428***	-0.197	-0.298**	-0.504***	-0.307*	-0.582***
	(0.117)	(0.101)	(0.092)	(0.084)	(0.138)	(0.060)
Farmer	-0.367**	-0.049	0.076	0.052	0.508***	0.354***
	(0.118)	(0.106)	(0.097)	(0.089)	(0.129)	(0.070)
R-squared	0.008	0.006	0.006	0.007	0.010	0.011
Observations	129,213	150,622	172,001	209,366	89,466	471,491
Lifespan after 60						
SES (ref: Skilled manual)						
White collar	0.091	0.206*	0.298***	0.611***	1.064***	1.053***
	(0.118)	(0.098)	(0.087)	(0.073)	(0.112)	(0.043)
Low-skilled manual	-0.013	0.134	0.089	-0.126	-0.043	-0.284***

	(0.098)	(0.085)	(0.075)	(0.064)	(0.099)	(0.043)
Unskilled manual	-0.229*	-0.210*	-0.181*	-0.377***	-0.302*	-0.525***
	(0.099)	(0.086)	(0.078)	(0.072)	(0.120)	(0.052)
Farmer	-0.291**	-0.050	0.084	0.064	0.393***	0.276***
	(0.100)	(0.090)	(0.083)	(0.076)	(0.112)	(0.061)
R-squared	0.007	0.007	0.005	0.008	0.010	0.011
Observations	114,988	134,637	153,869	188,418	83,248	449,490
Lifespan after 70						
SES (ref: Skilled manual)						
White collar	0.103	0.125	0.383***	0.539***	0.927***	0.863***
	(0.099)	(0.084)	(0.074)	(0.062)	(0.096)	(0.036)
Low-skilled manual	-0.017	0.073	0.079	-0.111*	0.016	-0.214***
	(0.082)	(0.071)	(0.064)	(0.055)	(0.084)	(0.036)
Unskilled manual	-0.200*	-0.098	-0.109	-0.266***	-0.131	-0.428***
	(0.083)	(0.073)	(0.067)	(0.061)	(0.103)	(0.044)
Farmer	-0.164	-0.039	0.085	0.029	0.250**	0.118*
	(0.084)	(0.076)	(0.071)	(0.064)	(0.096)	(0.052)
R-squared	0.008	0.006	0.004	0.009	0.011	0.012
Observations	90,060	105,601	120,516	151,060	70,869	403,949
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses.

Sources: See Table 1

Table B2. Estimates from accelerated failure time model of remaining lifespan at age 40, 1861-1920 cohorts.

	1841-1850 (1)	1851-1860 (2)	1861-1870 (3)	1871-1870 (4)	1891-1900 (5)	1911-1920 (6)
Panel A. Men						
SES (ref: Skilled manual)						
White collar	-0.024**	-0.032***	-0.031***	-0.038***	-0.006	0.019***
	(0.009)	(0.008)	(0.007)	(0.005)	(0.008)	(0.002)
Low-skilled manual	0.029***	0.032***	0.008	0.003	0.035***	-0.014***
	(0.008)	(0.007)	(0.006)	(0.005)	(0.006)	(0.002)
Unskilled manual	0.004	0.005	-0.005	-0.001	0.038***	-0.013***
	(0.007)	(0.006)	(0.006)	(0.005)	(0.007)	(0.002)
Farmer	0.045***	0.044***	0.030***	0.026***	0.061***	0.034***
	(0.008)	(0.007)	(0.006)	(0.005)	(0.007)	(0.003)
R-squared	0.022	0.021	0.017	0.013	0.009	0.009
Observations	146,921	166,918	189,844	237,526	103,362	505,997
Panel B. Women						
SES (ref: Skilled manual)						
White collar	0.035***	0.040***	0.029***	0.032***	0.048***	0.034***
	(0.009)	(0.007)	(0.006)	(0.005)	(0.007)	(0.002)
Low-skilled manual	0.010	0.015*	0.002	-0.006	-0.001	-0.011***
	(0.008)	(0.007)	(0.006)	(0.005)	(0.006)	(0.002)
Unskilled manual	-0.020*	-0.005	-0.025***	-0.031***	-0.015*	-0.016***
	(0.008)	(0.007)	(0.006)	(0.005)	(0.007)	(0.003)
Farmer	-0.008	0.007	0.005	0.000	0.018*	0.014***
	(0.008)	(0.007)	(0.006)	(0.006)	(0.007)	(0.003)
R-squared	0.007	0.004	0.003	0.004	0.006	0.006
Observations	139,785	162,501	185,108	224,249	93,443	482,188
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses.

Sources: See Table 1.

Table B3. LPM estimates of probability of survival until ages 50, 60 and 70, 1861-1920 cohorts.

	1841-1850 (1)	1851-1860 (2)	1861-1870 (3)	1871-1870 (4)	1891-1900 (5)	1911-1920 (6)
Panel A: Men						
P(survival past age 50)						
SES (ref: Skilled manual)						
White collar	-0.004	-0.001	-0.003	-0.006**	-0.002	-0.001
	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)
Low-skilled manual	0.005	0.008**	0.001	-0.002	0.005*	-0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Unskilled manual	0.000	0.000	-0.003	-0.004*	0.006*	-0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Farmer	0.010***	0.009***	0.006*	0.004	0.008**	0.004***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
R-squared	0.011	0.009	0.007	0.004	0.002	0.001
Observations	146,921	166,918	189,844	237,526	103,362	505,997
P(survival past age 60)						
SES (ref: Skilled manual)						
White collar	-0.017***	-0.020***	-0.019***	-0.023***	-0.011**	0.001
	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.001)
Low-skilled manual	0.014**	0.015***	0.005	0.004	0.016***	-0.005***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.001)
Unskilled manual	0.000	-0.001	-0.005	0.001	0.015***	-0.004**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)
Farmer	0.020***	0.022***	0.016***	0.012***	0.024***	0.015***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)
R-squared	0.019	0.018	0.013	0.010	0.006	0.004
Observations	146,921	166,918	189,844	237,526	103,362	505,997
P(survival past age 70)						
SES (ref: Skilled manual)						
White collar	-0.028***	-0.039***	-0.042***	-0.040***	-0.014*	0.015***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.006)	(0.002)
Low-skilled manual	0.020***	0.020***	0.010**	0.011***	0.028***	-0.012***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.005)	(0.002)
Unskilled manual	0.005	0.006	0.004	0.010**	0.032***	-0.009***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.005)	(0.002)
Farmer	0.029***	0.031***	0.023***	0.023***	0.050***	0.031***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)
R-squared	0.021	0.022	0.019	0.016	0.011	0.008
Observations	146,921	166,918	189,844	237,526	103,362	505,997
Panel B: Women						
P(survival past age 50)						
SES (ref: Skilled manual)						
White collar	0.012***	0.013***	0.009***	0.007***	0.005*	0.003***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Low-skilled manual	0.004	0.002	-0.002	-0.002	-0.000	-0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Unskilled manual	-0.003	-0.001	-0.007**	-0.007***	-0.002	-0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)
Farmer	0.001	0.003	0.001	-0.001	0.002	0.002*
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
R-squared	0.005	0.002	0.001	0.002	0.001	0.001
Observations	139,785	162,501	185,108	224,249	93,443	482,188
P(survival past age 60)						

SES (ref: Skilled manual)						
White collar	0.012*	0.015***	0.009**	0.009**	0.018***	0.007***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.001)
Low-skilled manual	0.001	0.005	0.000	-0.002	0.002	-0.004**
	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)
Unskilled manual	-0.013**	-0.002	-0.012***	-0.013***	-0.002	-0.003*
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.001)
Farmer	-0.004	0.003	0.001	-0.002	0.007*	0.005**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)
R-squared	0.006	0.003	0.003	0.002	0.002	0.002
Observations	139,785	162,501	185,108	224,249	93,443	482,188
P(survival past age 70)						
SES (ref: Skilled manual)						
White collar	0.009	0.017***	0.007	0.016***	0.028***	0.020***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)	(0.002)
Low-skilled manual	-0.000	0.008	0.001	-0.004	-0.002	-0.008***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)
Unskilled manual	-0.015**	-0.008	-0.015***	-0.019***	-0.012*	-0.010***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)
Farmer	-0.012*	0.002	0.001	0.001	0.015**	0.012***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.002)
R-squared	0.005	0.005	0.004	0.004	0.005	0.004
Observations	139,785	162,501	185,108	224,249	93,443	482,188
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses.

Sources: See Table 1.

Table B4. OLS estimates of remaining lifespan at age 40 for women, 1861-1920 cohorts.

	1841-1850 (1)	1851-1860 (2)	1861-1870 (3)	1871-1870 (4)	1891-1900 (5)	1911-1920 (6)
Panel A: SES based on own occupation						
SES (ref: Skilled manual)						
White-collar	-0.615	0.097	0.167	0.936***	2.297***	2.289***
	(0.527)	(0.394)	(0.290)	(0.212)	(0.244)	(0.131)
Low-skilled manual	-0.583	-0.239	-0.059	0.094	0.579**	0.620***
	(0.430)	(0.344)	(0.258)	(0.192)	(0.218)	(0.129)
Unskilled manual	-3.237***	-1.895***	-2.228***	-1.769***	-0.315	-0.194
	(0.517)	(0.426)	(0.328)	(0.254)	(0.362)	(0.158)
Farmer	-0.613	-1.127*	-0.827	-0.000	0.138	0.483
	(0.678)	(0.549)	(0.505)	(0.451)	(0.493)	(0.356)
R-squared	0.007	0.008	0.006	0.009	0.017	0.012
Observations	26,230	31,981	38,691	53,354	27,359	135,335
Panel B: SES based on spouse's occupation						
SES (ref: Skilled manual)						
White-collar	0.606***	0.749***	0.624***	0.746***	1.022***	1.045***
	(0.173)	(0.147)	(0.130)	(0.112)	(0.177)	(0.055)
Low-skilled manual	0.013	0.200	-0.055	-0.332***	-0.375*	-0.578***
	(0.148)	(0.128)	(0.115)	(0.100)	(0.159)	(0.057)
Unskilled manual	-0.253	-0.070	-0.272*	-0.491***	-0.578**	-0.608***
	(0.145)	(0.125)	(0.114)	(0.106)	(0.176)	(0.068)
Farmer	-0.268	0.052	0.106	-0.076	0.219	0.316***
	(0.146)	(0.131)	(0.120)	(0.110)	(0.166)	(0.079)
R-squared	0.009	0.007	0.006	0.007	0.010	0.009
Observations	113,789	131,027	147,053	173,057	70,273	400,600
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses.

Sources: See Table 1.

Table B5. OLS estimates of remaining lifespan at age 40, 1861-1900 cohorts. Sample of links based on exact matches of two first names.

	1841-1850		1851-1860		1861-1870		1871-1870		1891-1900	
	(1)	(2)	(3)	(4)	(5)	(6	(7)	(8)	(9)	(10)
Panel A: Men										
SES (ref: Skilled manual)										
White collar	-0.740***	-0.572*	-1.005***	-0.886***	-0.885***	-0.850***	-0.906***	-0.905***	-0.075	0.075
	(0.150)	(0.226)	(0.130)	(0.186)	(0.113)	(0.151)	(0.095)	(0.121)	(0.148)	(0.172)
Low-skilled manual	0.587***	0.675**	0.661***	0.631***	0.260*	0.270*	0.238**	0.263*	0.838***	0.771***
	(0.132)	(0.208)	(0.115)	(0.170)	(0.102)	(0.138)	(0.085)	(0.108)	(0.126)	(0.144)
Unskilled manual	0.108	0.222	0.090	-0.047	-0.002	0.016	0.117	0.151	0.853***	0.778***
	(0.125)	(0.196)	(0.109)	(0.161)	(0.098)	(0.133)	(0.085)	(0.109)	(0.129)	(0.148)
Farmer	0.868***	0.932***	0.898***	0.682***	0.711***	0.746***	0.680***	0.710***	1.592***	1.553***
	(0.130)	(0.209)	(0.118)	(0.178)	(0.108)	(0.151)	(0.096)	(0.125)	(0.137)	(0.156)
R-squared	0.031	0.037	0.031	0.036	0.026	0.028	0.021	0.021	0.015	0.015
Observations	146,921	60,149	166,918	78,877	189,844	104,763	237,526	147,455	103,362	78,493
Panel A: Women										
SES (ref: Skilled manual)										
White collar	0.494**	0.713**	0.662***	0.649**	0.561***	0.519**	0.788***	0.804***	1.470***	1.389***
	(0.162)	(0.274)	(0.135)	(0.209)	(0.116)	(0.158)	(0.098)	(0.123)	(0.145)	(0.165)
Low-skilled manual	0.056	-0.086	0.256*	0.180	0.062	-0.046	-0.179*	-0.214	-0.011	0.024
	(0.136)	(0.238)	(0.117)	(0.185)	(0.102)	(0.141)	(0.087)	(0.110)	(0.128)	(0.143)
Unskilled manual	-0.490***	-0.653**	-0.215	-0.289	-0.489***	-0.582***	-0.697***	-0.642***	-0.378*	-0.288
	(0.138)	(0.241)	(0.119)	(0.189)	(0.107)	(0.148)	(0.097)	(0.123)	(0.156)	(0.176)
Farmer	-0.301*	-0.478*	0.059	-0.108	0.106	-0.018	0.011	0.049	0.549***	0.557***
	(0.139)	(0.243)	(0.124)	(0.199)	(0.113)	(0.158)	(0.103)	(0.132)	(0.146)	(0.164)
R-squared	0.009	0.011	0.006	0.006	0.006	0.007	0.007	0.007	0.010	0.010
Observations	139,785	49,973	162,501	66,279	185,108	99,042	224,249	141,391	93,443	74,416
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Restricted	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses.

Sources: See Table 1.

CENTRE FOR ECONOMIC DEMOGRAPHY

The Centre for Economic Demography (CED), established in 2006 as a Linnaeus Centre of Excellence, is a multidisciplinary research centre at Lund University with members from the School of Economics and Management, the Faculty of Medicine and the Faculty of Social Sciences. The mission of CED is to promote research and training in population studies. Researchers at the Centre study demographic issues from different perspectives and in a variety of contexts across time and space. More information at www.ed.lu.