

**Socioeconomic Status and Adult Life Expectancy 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 7083
20007 Lund, Sweden

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

Acknowledgements

This paper has been done within the research project “The Rise and Fall of the Industrial City”, funded by the Swedish Foundation for the Humanities and Social Sciences (RJ). It also forms part of the LONGPOP project funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676060. (disclaimer: this publication reflects only the author’s view and that the Agency is not responsible for any use that may be made of the information it contains). Eriksson furthermore acknowledge 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), Southern Danish University Odense, Umeå University, Stockholm University, Uppsala University, SSB (Oslo), EHS (Keele) and IFN (Stockholm).

Abstract

Across the developed world a socioeconomic gradient in health and adult mortality exists. Our knowledge about conditions in the more distant past is much more rudimentary: we do not know exactly when and why the mortality gradient emerged, or whether health inequalities is a salient historical fact. In this paper we study differences in life expectancy at age 40 by socioeconomic status. The analysis is based on individual-level mortality data covering the entire population of Sweden, which have been linked to the full count censuses of 1880, 1890, 1900, 1910, 1930 and 1950 using probabilistic linking methods. This allows us to control for selection through unobserved spatial and family-of-origin heterogeneity. Our findings show that initially the life expectancy of upper/middle class men were shorter than that of the working class, while farmers had the longest life expectancy of all groups. The modern positive gradient between socioeconomic status and life expectancy does not become evident until the 1950's. For women the modern gradient is apparent in all cohorts, with the longest life expectancy for the high-status groups, and the shortest for low-status groups, although differences were smaller. These results suggest that today's steep mortality gradient is of a recent origin which, counterintuitively, coincided with the development of modern medicine and the welfare state.

Keywords: Socioeconomic status, Inequality, Mortality, Life expectancy

The positive correlation between socioeconomic status (SES) and adult life expectancy in modern developed countries is a well-established fact. Whether measured by income, education or occupational status, SES is positively associated with health and negatively associated with all-cause mortality (e.g. Elo 2009; Mackenbach et al. 2003; Smith 1999, 2004; Marmot 2004; Torssander and Erikson 2010). Recent estimates of the association between income and life expectancy puts the issue into perspective: Chetty et al (2016) find that the 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. Comparable inequalities are also present in more egalitarian settings: Swedish men and women belonging to the top income quintile are expected to live for 9.1 and 7.1 years longer after age 35 compared to men and women in the bottom quintile (Hederos et al 2018). 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 related to smoking-related cancers and cardiovascular diseases (Hederos et al. 2017; Mackenbach et al. 2015)¹.

While health inequalities in contemporary societies is 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 today's (e.g. Antonovsky 1967; Smith 2004). In a recent review, Elo (2009) state 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 have always been able to avoid premature deaths regardless of the specific mechanisms and diseases dominating mortality. The empirical support for the historical validity of these claims is, however, based on either aggregate statistics or limited individual level samples which are not directly comparable to modern estimates, and when

¹ Also see Bronnum-Hansen and Baadsgaard 2007; Burström, K. et al. 2005; Kunst et al. 2004; Mackenbach et al. 2003; Shkolnikov et al. 2012; Statistics Sweden 2016; Steingrimsdottir et al. 2012. In addition, the variance in life span is greater for low-status groups than for high-status groups, and has been reduced more over time for the high-status groups than for the low-status ones (Van Raalte et al. 2014).

scrutinized further, far from unanimous (see Bengtsson and Van Poppel 2011 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-WWII period (Vågerö and Norell 1989; Kunst et al. 1990; Bengtsson, et al. 2017).

The aim of this paper is to provide the first population based individual-level estimates of the SES gradient in mortality before World War II. By combining micro-level census data and mortality 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 mortality records of all decedents in Sweden between 1830 and 2020 from the Swedish Death Index (2021). 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, a pressing issue given the high levels of child and in particular infant mortality that prevailed in the 19th century. 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 life expectancy. 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 life expectancy for 1841-1900 male cohorts was negative and the opposite of the modern gradient: white-collar men had the lowest life expectancy of all groups of men and women. Only the last cohorts which we study, born 1911-1920, display the modern positive gradient between SES and life expectancy. The results refute the presumption that there is a consistent negative correlation between SES and mortality. For women we find a positive relationship between SES and life expectancy for all cohorts born 1841-1920, similar to the pattern prevailing today. In terms of magnitudes, we observe historical inequalities in adult life expectancy that are much smaller than the large differences found today between the top and bottom of the income distribution: differences between the longest and shortest living groups rarely exceed 3 years and are in most cases much smaller.

Taken together, our results clearly show that the strong positive relationship between SES and mortality is of a recent origin and not a salient historical fact, and that, counterintuitively, the relationship between SES and mortality emerged and strengthened parallel to the development of modern medicine and the expansion of the welfare state.

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 outlined by Omran (1971) Link and Phelan (1995; 1996; 2002; 2004) identify SES as fundamental cause of mortality, favouring 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; 1980) 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 a severe reduction in the prevalence of the disease and a 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 for whether SES was an important determinant of 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 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 Razzle and Spence (2006) whom, 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.

Micro level data based on geographically limited samples addresses some of the shortcomings of aggregate statistics but suffers from selection, a lack of generalizability and small numbers. An early study by Chapin (1924) of mortality differences in Providence, Rhode Island, in 1865, found that non-taxpayers had higher overall mortality than taxpayers.²

² In addition to overall differences, Chapin (1924) also find 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.

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 according to SES, and mixed evidence for the correlation between wealth and mortality in a rural sample based on the 1850 and 1860 U.S. federal censuses. For Sweden the available evidence seems to show that a clear negative gradient between SES and mortality only emerged sometime after the Second World War (Bengtsson and Dribe 2011; Bengtsson, Dribe and Helgertz 2017; Edvinsson and Broström 2012).

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 into one of the world's richest and most equal industrialized countries (Schön 2000). Although the full development of the welfare state took place later, some 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.

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 beginning in the 1930s (Bengtsson and Ohlsson 1994; Hofsten and Lundström 1976).

The transition is illustrated in Figure 1 which plots period life expectancy at age 40 (e_{40}) from 1751 to 2017. 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 2017 ($e_{40} = 44.8$ for women and 41.9 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 Dribe 2005; Bengtsson and Ohlsson 1985). The second phase lasted from the beginning of the nineteenth century until the beginning of the twentieth. Although mortality was still subject to significant yearly variation, it was on a much more limited scale 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 yearly variation and epidemic outbreaks. In its place heart disease and cancer has emerged as the dominant causes of death.

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 sourced from the 1880, 1890, 1900, 1910, 1930 and 1950 censuses³. 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 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 etc. Our main outcome variable is life expectancy

³The 1880-1910 full-count censuses have been digitized by the Swedish National Archives and published by the North Atlantic Population Project (NAPP, www.nappdata.org). The 1930 census 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.

at age 40 (e_{40}) which is the earliest age at which life expectancy is measurable in a consistent manner for all cohorts in the sample.

B. Occupational coding and class scheme

We use the HISCLASS-scheme to measure SES (Van Leeuwen and Maas 2011). HISCLASS is based on the HISCO coding scheme (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. In the main specification we aggregate HISCLASS into five more general groups: 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 (HISCLASS 8). 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 wives⁴. 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. This makes the linking of individuals between the 1950 census and the death index straightforward. Because the 1880-1930 censuses

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.

precede the introduction of modern identification numbers, we rely on probabilistic linking methods for identifying individuals in each source. We use birthplace (parish of birth), sex and birthyear as index variables, meaning that individuals are only considered possible matches if these variables are identical in the death index and 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. The 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).

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 (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 (Eriksson 2015; Wisselgren et al. 2014) and exceeds rates achieved for North American and British censuses (Ferrie and Long 2013).

Overall, the linked samples resemble the censuses apart from some minor differences (Appendix: table A2). The shares of white-collar, manual skilled and low-skilled status men and women in the linked sample are near identical to the distributions observed in the censuses. The share of farmers is higher in the linked sample 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 life expectancies for the earlier cohorts calculated from the linked samples of men and women is somewhat higher than the life expectancies calculated using the complete death index or estimates provided by the Human Mortality Database (Appendix: figure A1). The series do however track each other closely in terms of trends and yearly deviations.

D. Sample selection

We limit our analysis to cohorts born between 1841-1920 observed in the censuses when aged 30-39. 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 numbers to 1,510,325 men and 1,531,540 women. We exclude men with no recorded occupation in the census, and women with no own or spouse occupation, and all 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. Table 1 displays the descriptive statistics of the analytical sample for men and women.

Table 1

IV. Estimation

Our primary outcome of interest is life expectancy at age 40. Because the sample only includes extinct cohorts, observations are not subject to censoring. Cohort life expectancy therefore equals the arithmetic mean of ages at death (the total person years lived in the cohort divided by the number of individuals in the cohort). In order to account for a number of confounders, the relationship between SES and life expectancy at age 40 is estimated using the following linear fixed-effects regression model⁶:

$$e_{40i} = \alpha + \beta SES_i + \beta Cohort_i + \beta SES_i Cohort_i + \beta X_i + \partial_j + \epsilon_{ity} \quad (1)$$

In which e_{40ij} is life expectancy at age 40 for individual I , SES_i is the socioeconomic status of the individual, $Cohort_i$ denotes 5-year birth 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 change 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

⁶ See de la Croix and Licandro 2015, and Cummins 2017 for similar approaches to estimating adult life expectancy.

water, organization of poor relief and health care, breast-feeding practices, and differences in agricultural productivity (Edvinsson and Lindkvist 2011; Garrett et al 2001; Reid 1997; Smith 1983; van Poppel et al. 2005; Woods et al. 1993). 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 ($\hat{\theta}_j$) of fixed effects 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 life expectancy for the different SES groups (see Table 2). White-collar men belonging to the earliest observed cohort (1841-1845) had the shortest life expectancy 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 life expectancy of white collar now exceeds the manual groups by at least a year, and only trails farmer by 6 months. For women, differences between groups are much less pronounced, 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. Differences does however become more pronounce for later cohorts and has almost doubled 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 cohorts born in the middle of the 19th century, but did not appear for men until the beginning of the 20th century.

Table 2

B. Main results

Figure 3 presents the results of estimations of our main model specified in (1) in the form of predicted life expectancy for men and women. The basic pattern observed in Table 2 remains after controlling for both demographic and geographic confounders. Overall, absolute differences in life expectancy remained stable for men born 1861-1900, with all groups following a common trend of increasing life expectancy. 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 life expectancy 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 life expectancy from the highest status to the lowest is apparent for all cohorts. Women in the white-collar group consistently have the longest life expectancy at 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 life expectancy between the top and bottom.

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

C. Selection on family background

How much of the estimated differences in life expectancy 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). 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 et al 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 19th century Swedish infant and child mortality rates: low status groups had significantly higher rates of infant mortality compared to high status groups (Bengtsson 2004; Edvinsson et al 2005; Molitoris and Dribe 2016). These differences in childhood conditions implies that bias becomes more severe as SES falls, and to be positive if selection, or negative if scarring, is the dominant mechanism. We address the issue by considering family background and unobserved heterogeneity shared between siblings. By 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 may be observed in preceding census.

We address selection by first examining the relationship between father SES and life expectancy independently, and whether our estimates of the association between own SES and life expectancy 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.

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.

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 life expectancy mirrors estimates for own SES. When controlling for both own and father SES the gradients remain across both aspects of SES. Importantly, for men, the shift in the gradient across cohorts from negative to positive the negative SES gradient for earlier cohorts SES 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 life expectancy 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 life expectancy for both men and women remains 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 life expectancy. That the sibling fixed-effects model yield 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 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 life expectancy. 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 variable to the model, and even remain after estimating sibling fixed effects models.

Figure 4

The results using the occupational income scores across all censuses and cohorts offers 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 life expectancy for high income (the 4th quartile group) 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

How generalizable are our results across different 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 fit separate models separately for every Swedish county, yielding 24 county specific estimates of the relationship between SES and mortality for men and women by cohort. The regression coefficients for SES (with Skilled manual set as the reference category) are presented in Figure 5. Overall, for both men and women, the direction of the county specific gradients conforms to the general results established prior. 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 life expectancy 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 penalties and premiums to life expectancy 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 19th and the beginning of the 20th 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 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 life expectancy at more advanced ages we also use e_{50} , e_{60} and e_{70} as alternative outcome variables (see 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 all 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 extended life expectancy, although not as much as farmers; a result indicating that either living on a farm or working with farming had positive implications for life expectancy. 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 life expectancy and the lowest-status group, unskilled workers, have the shortest. Moreover, within the white-collar group there are differences corresponding to social status. Similarly, among non-farm manual workers mortality vary by skill-level.

In our main model, women's SES is based on own occupations, or in its absence, the spouses occupation. In Table B4 estimates are presented separately for women with SES based on own occupation, and SES based on spouse 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 only include 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.

VI. Discussion

The results presented in this paper demonstrate the existence of a negative male and positive female relationship between SES and life expectancy at the end of the 19th and beginning of the 20th century in Sweden. For women the pattern resembles the modern SES gradient with white-collar women having the longest life expectancy, although differences between groups were relatively small. Among men, farmers experienced the longest life after 40 initially, followed by manual workers, while white-collar men had the shortest life expectancy. A similar observation has been made for the northern town of Sundsvall in the same period but was noted

as a paradox and difficult to explain (Edvinsson 1992). This raises the question of why white-collar men, the most privileged group with regards to status and resources, were also the most disadvantaged group in terms of mortality?

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 death index does not include information about causes of death⁸, it is not possible to estimate SES differences in cause specific mortality for our linked individual level sample. We therefore turn to aggregate statistics on male cause specific mortality for different occupational groups published by Statistics Sweden (1911; 1921; 1931). Proportionate cause specific mortality for men is presented in Table 8. The occupational groups to a large extent 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 consists of manual workers, but also includes a small number of managers, owners and administrators which would fall into the white-collar group if classified according to the HISCLASS scheme. The public service group is more narrowly defined and closely resembles the white-collar group, consisting of priests, teachers, doctors, and civil servants.

Table 7

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 strokes). 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

⁸ The Swedish Death Index is a commercial genealogical resource. Therefore, due to confidentiality, causes of death were never collected from the original sources.

the group most affected by circulatory diseases⁹. Public service workers were thus forerunners in the sense that cause-specific mortality by the end of the 19th 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 for why low SES is related to poor health and higher mortality (Adler and Stewart 2010; Cavelaars et al. 2000; Elo 2009; Marmot 2004; Smith 1999). 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; Rodu and Cole 2004). It was not until after World War II, when cigarettes became increasingly popular, that social and sex 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 (Nordlund 2005).

Because the risks associated with tobacco consumption were still unknown, the public debate and policy during the 19th century primarily focused on the effects of alcohol, and the

⁹ This result is similar Logan (1952) whom show 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.

consumption of the lower classes. 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 sex, marital status and social standing served to further limit the alcohol consumption of women and the working class (Brunn and Frånberg 1985).

Although relatively few people were overweight in 19th 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 (Nilsson 2011; Bildtgard 2002) 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). Table 9 summarizes the yearly consumption by household income. 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. These differences are significant and consistent with the mortality patterns observed among the public service occupations in Table 8.

Table 8

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 adult life expectancy 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 20th 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. This study deals with SES differences in adult life expectancy in a period before the full development of the welfare state, when large segments of society still lived in what we would today consider deep poverty, while the privileged enjoyed considerable affluence. It was a context where one could perhaps expect pronounced SES differences in mortality. This was also the case for children in Stockholm, the capital of Sweden, who in this period suffered much higher mortality if they belonged to the lower classes than if they were born into the upper or middle classes (Burström and Bernhardt 2001; Burström, B. et al. 2005; Molitoris and Dribe 2016). For adults, however, our findings suggest a very different picture.

Some of the observed difference in life expectancy can be explained by observed characteristics apart from SES, and even more by unobserved parish-level factors. Still, about half of the crude difference in life expectancy by SES remains after fully controlling for these factors. Our findings clearly show that adult life expectancy 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 life spans 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 adult life expectancy. A similar observation has been made for the northern town of Sundsvall in the same period but was noted as a paradox which was difficult to explain (Edvinsson 1992: 191; see also Razzel and Spence 2006 for a similar argument for Britain). For women we instead found a pattern resembling the modern SES gradient with the shortest life expectancy in the working class and the highest among white-collar groups. 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 life style, was gaining in prominence. White-collar men stand out as a group whom had the means to consume richer and unhealthier food, 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 life expectancy. This conclusion is also 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 life expectancy compared to other classes.

The longer life expectancy for 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. We do not know the extent to which consumption of smoking tobacco or alcohol differed between farmers and blue-collar workers, but it 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, wives of unskilled workers had the shortest life expectancy 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. Naturally, this is mere speculation, and we need more evidence on actual living conditions and cause-of-death specific mortality across SES groups 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, however, 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 life span become an established fact for both men and women.

References

- 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.
- Bengtsson, T., & Dribe, M. (2005). New evidence on the standard of living in Sweden during the 18th and 19th centuries: Long-term development of the demographic response to short-term economic stress. In R. C. Allen, T. Bengtsson & M. Dribe (Eds.), *Living Standards in the Past. New Perspectives on Well-being in Asia and Europe* (pp. 341–372). Oxford: Oxford University Press.
- 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., & Ohlsson, R. (1994). The demographic transition revised. In T. Bengtsson (Ed.), *Population, Economy, and Welfare in Sweden* (pp. 13–35). Berlin-Heidelberg: Springer Verlag.
- 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. (2017). Income, occupational status, and mortality in the long run: southern Sweden 1815–2011. Paper for the annual meeting of the Population Association of America, Chicago April 2017.
- 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.

- Bronnum-Hansen, H., & Baadsgaard, M. (2007). Increasing social inequality in life expectancy in Denmark. *European Journal of Public Health, 17*, 585–586.
- 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.
- Burström, K., Johanneson, M., & Diderichsen, F. (2005). Increasing socio-economic inequalities in life expectancy and QALYs in Sweden 1980–1997. *Health Economics, 14*, 831–850.
- Burström B., Macassa G., Öberg L., Bernhardt E., & Smedman L. (2005). Public health then and now. Equitable child health interventions: the impact of improved water and sanitation on inequalities in child mortality in Stockholm, 1878–1925. *American Journal of Public Health, 95*, 208–216.
- 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.
- Chapin, C. V. (1924). Deaths among taxpayers and non-taxpayers, income tax, Providence, 1865. *American Journal of Public Health, 14*, 647–651.
- 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.
- Easterlin, R. A. 1996. How beneficent is the market? A look at the modern history of mortality. *European Review of Economic History, 3*, 257–294

- Edvinsson, S. 1992. *Den osunda staden. Sociala skillnader i dödlighet i 1800-talets Sundsvall*. Stockholm: Almqvist & Wiksell International.
- 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.
- Eriksson, B. (2015). *Dynamic Decades. A Micro Perspective on Late Nineteenth Century Sweden* (PhD Dissertation). Lund: Lund Studies in Economic History 72.
- 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.
- Hederos, K., Jäntti, M., Lindahl, L., & Torssander, J. 2017. Trends in life expectancy by income and the role of specific causes of death. *Economica*, published online 28 Feb. 2017.
- Hofsten, E., & Lundström, H. 1976. *Swedish Population History: Main Trends from 1750 to 1970*. Stockholm: Liber Förlag.

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

Kinge J.M., Modalsli J.H., Øverland S. et al. (2019) Association of Household Income With Life Expectancy and Cause-Specific Mortality in Norway, 2005-2015. *JAMA*, 321:1916–1925.

Kunst, A. E., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., ... Mackenbach, J. P. (2004). Monitoring trends in socioeconomic inequalities in mortality: Experiences from a European Project. *Demographic Research Special Collection*, 2, 229–254.

Lambert, P. S., Zijdeman, R. L., Van Leeuwen, M. H. D., Maas, I., & Prandy, K. (2013). The construction of HISCAM: a stratification scale based on social interactions for historical comparative research. *Historical Methods*, 46, 77–89.

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

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.

Mackenbach, J.P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., ... Kunst, A.E. (2003). Widening socioeconomic inequalities in mortality in six Western European countries. *International Journal of Epidemiology*, 32, 830–837.

Mackenbach, J.P., Kulhánová, I., Bopp, M., Deboosere, P., Eikemo, T.A., Hoffmann, R., ... Lundberg, O. (2015). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the "fundamental causes" theory of social inequalities in health. *Social Science and Medicine*, 127, 51–62.

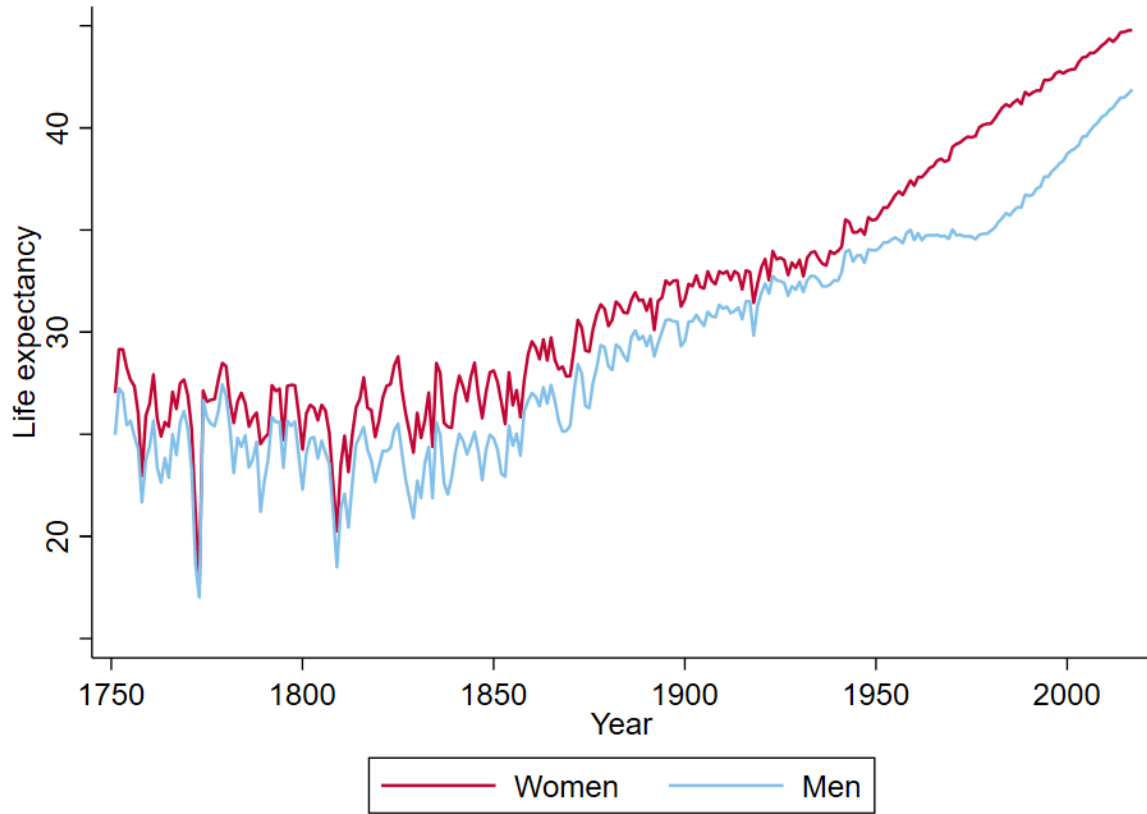
- Marmot, M. (2004). *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. London: Bloomsbury Press.
- Molitoris, J., & Dribe, M. (2016). Industrialization and inequality revisited: Mortality differentials and vulnerability to economic stress in Stockholm, 1878-1926. *European Review of Economic History*, 20, 176–197.
- Nordlund, A. (2005). Tobaksrökning och hälsa i Sverige under 1900-talet. In J. Sundin et al. (Eds.), *Svenska folkets hälsa i ett historiskt perspektiv* (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.
- Reid, A. (1997). Locality or class? Spatial and social differentials in infant and child mortality in England and Wales, 1895–1911. In C. Corsini & P-P. Viazzo (Eds.), *The Decline of Infant and Child Mortality: The European Experience* (129–154). Dordrecht: Martinus Nijhoff Publishers.
- Razzel, P., & Spence, C. (2006). The hazards of wealth: Adult mortality in pre-twentieth-century England. *Social History of Medicine* 19: 381–405.
- Rodu B., & Cole P. (2004). The burden of mortality from smoking: comparing Sweden with other countries in the European Union. *European Journal of Epidemiology*, 19, 129–131.
- Schön, L. (2000). *En modern svensk ekonomisk historia: tillväxt och omvandling under två sekel*. Stockholm: SNS.
- Shkolnikov, V.M., Andreev, E.M., Jdanov, D.A., Jasilionis, D., Kravdal, Ø., Vågerö, D., & Valkonen, T. (2012). Increasing absolute mortality disparities by education in Finland,

- Norway and Sweden, 1971–2000. *Journal of Epidemiology and Community Health*, 66, 372–378.
- Smith, D. S. (1983). Differential mortality in the United States before 1900. *Journal of Interdisciplinary History*, 13, 735–759.
- 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.
- Statistics Sweden (1915-43). *Dödsorsaker år 1911-1940*. Sveriges officiella statistik, folkmängden och dess förändringar.
- Statistics Sweden (2016). Livslängd och dödlighet i olika sociala grupper. *Demografiska rapporter 2016:2*. Stockholm: Statistics Sweden.
- Steingrimsdóttir, O.A., Næss, Ø., Moe, J.O., Grøholt, E.K., Thelle, D.S., Strand, B.H., Bævre, K. (2012). Trends in life expectancy by education in Norway 1961–2009. *European Journal of Epidemiology*, 27, 163–171.
- 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 Scheme*. 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.

- Van Raalte, A. A., Martikainen, P., & Myrskylä, M. (2014). Lifespan variation by occupational class: Compression or stagnation over time? *Demography*, *51*, 73–95.
- 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). Hälsa- och samhällsutveckling i Sverige 1750-2000. In J. Sundin et al. (eds.) *Svenska folkets hälsa i ett historiskt perspektiv* (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* (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.
- Woods, R., Williams, N., & Galley, C. (1993). Infant mortality in England 1550–1950. Problems in identification of long-term trends and geographical and social variations. In C. A. Corsini & P-P. Viazzo (Eds.), *The Decline of Infant Mortality in Europe 1800–1950. Four National Case Studies* (35–50). UNICEF and Istituto degli Innocenti di Firenze, Florence.

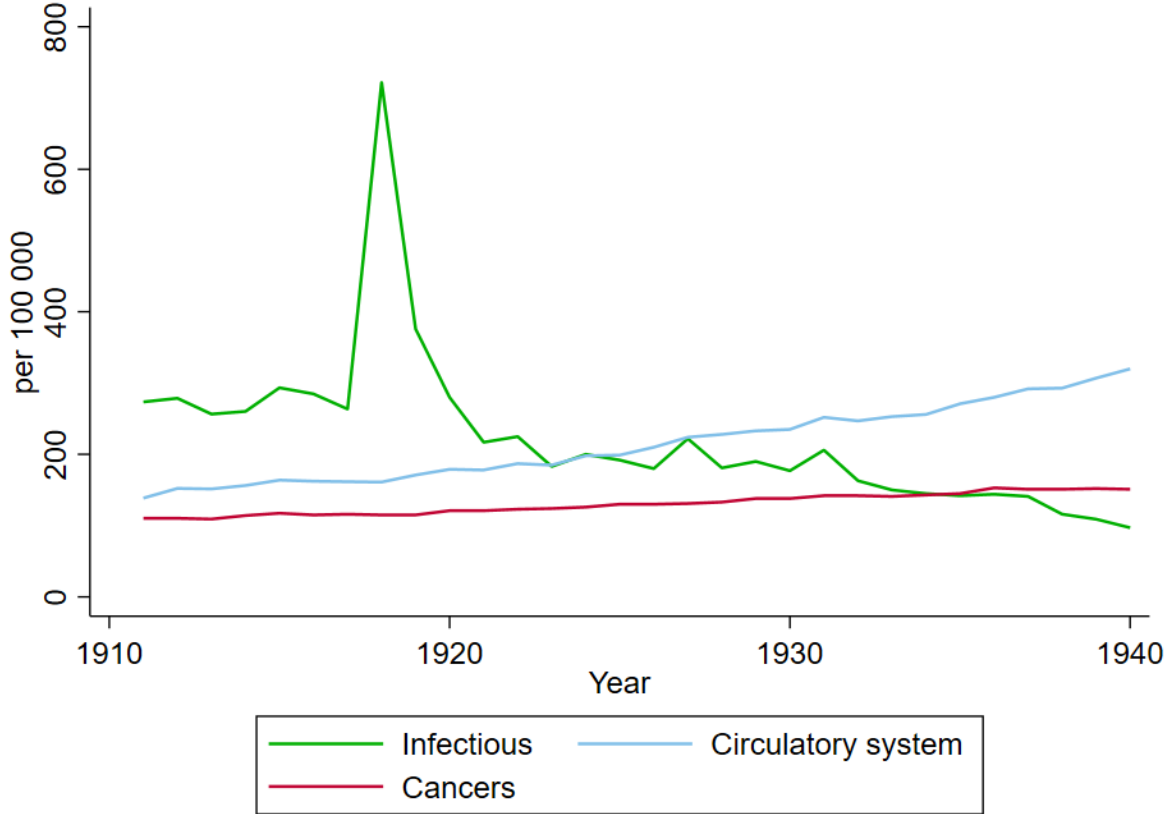
Figures

Figure 1. Period life expectancy at age 40 in Sweden 1751-2017.



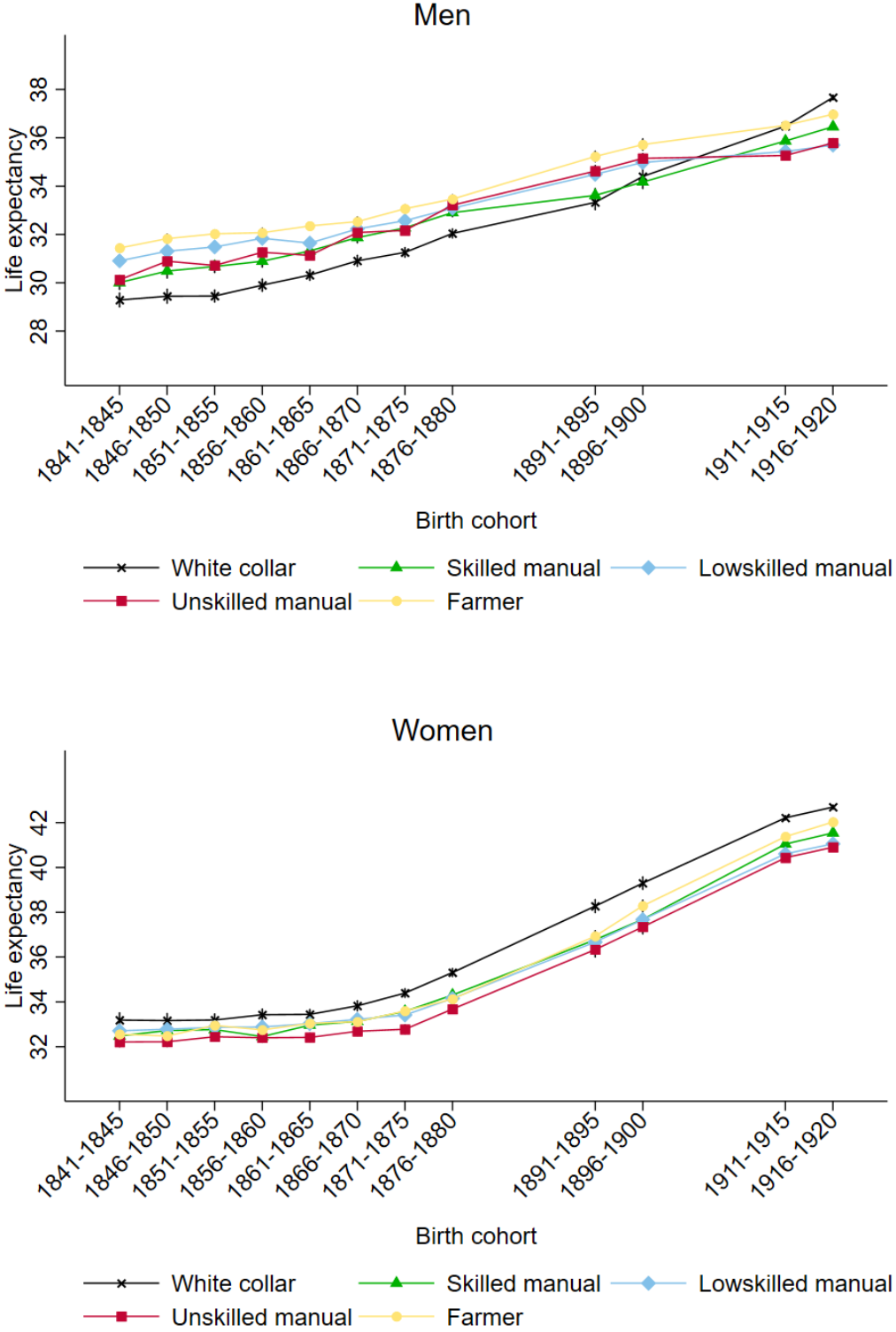
Source: Human Mortality Database

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



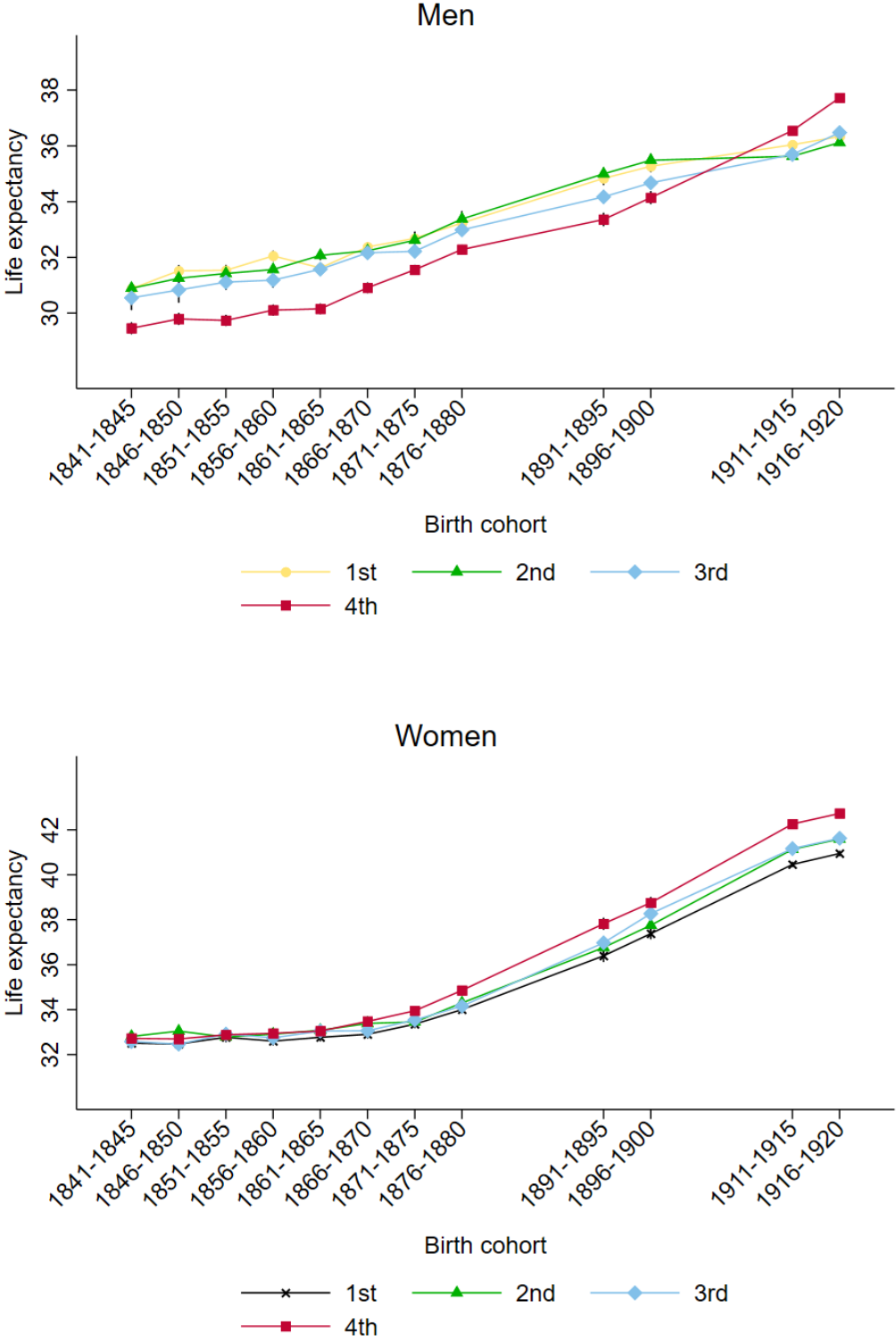
Source: Statistics Sweden (1911-1940)

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



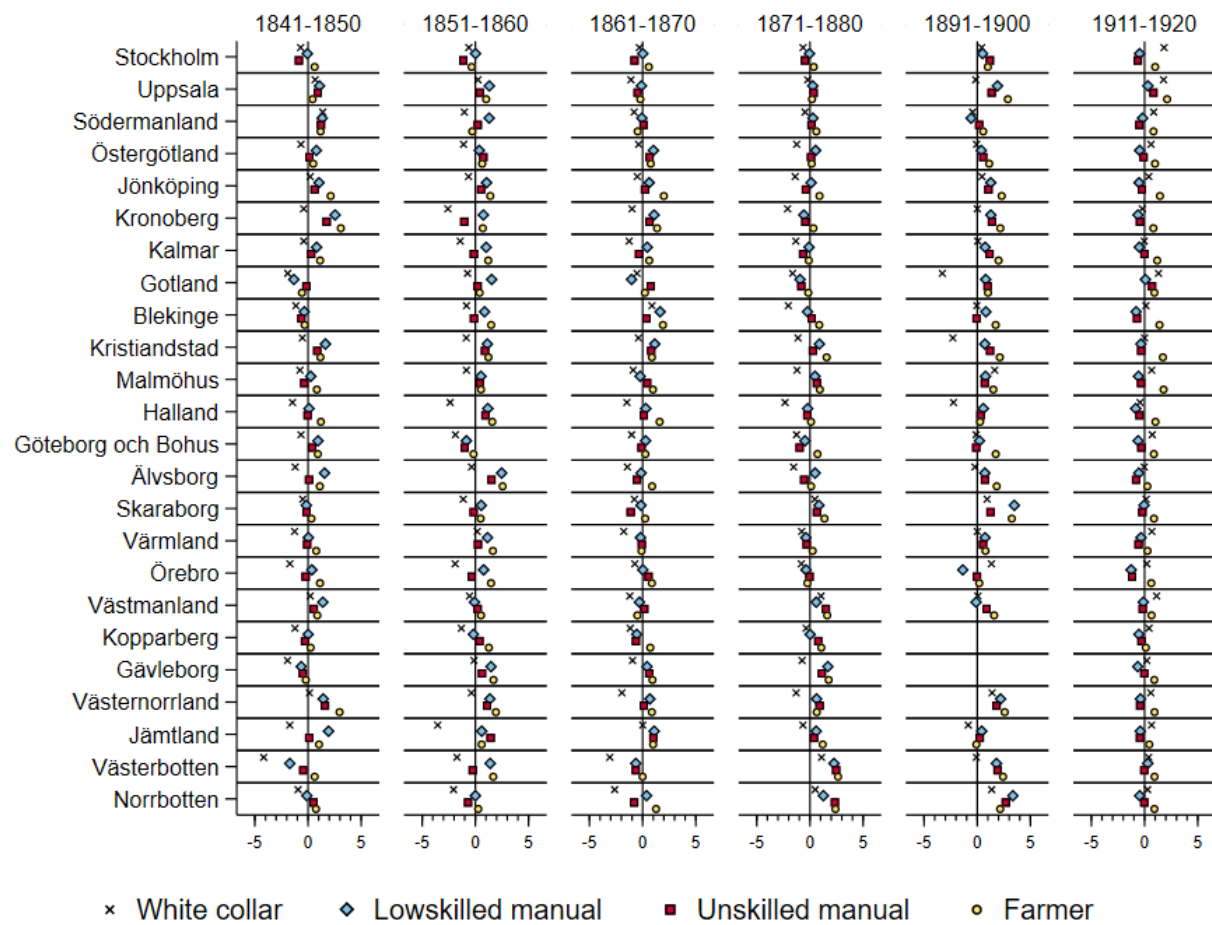
Note: The estimated models control for marital status, children, migrant status and includes fixed effects for county of birth and parish of residence.

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



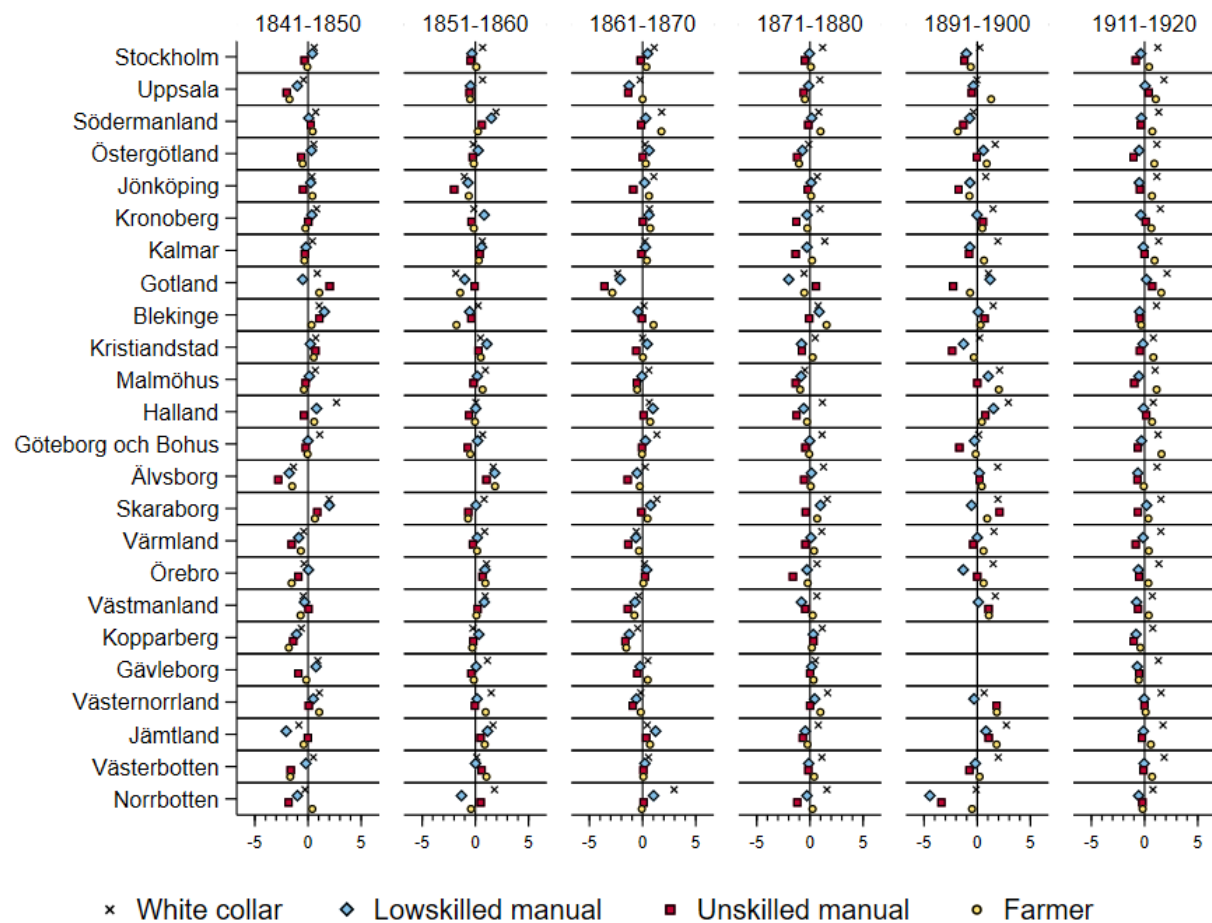
Note: See Figure 3.

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



Note: See Figure 3.

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



Note: See Figure 3.

Tables

Table 1. Descriptive statistics of analytical sample

	1880	1890	1900	1910	1930	1950	All
<i>Panel A. Men</i>							
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
e40	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
<i>Panel B. Women</i>							
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
e40	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

Table 2. Cohort life expectancy at age 40 by SES

Cohort	White collar	Skilled manual	Low-skilled manual	Unskilled manual	Farmer
<i>Panel A. Men</i>					
1841-1845	28.9	30	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	30.6	33
1856-1860	29.1	30.5	32.2	31.1	32.9
1861-1865	29.5	31	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
1896-1900	34	34	35	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
<i>Panel B. Women</i>					
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	33	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	33.7	34.2
1891-1895	38.2	36.8	36.6	36.3	37
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	40.9	42.2

Table 3. OLS estimates of life expectancy 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 (0.202)	0.0680 (0.213)		-0.325* (0.146)	0.0275 (0.152)		0.407* (0.199)	0.501* (0.203)		0.533*** (0.156)	0.352* (0.159)	
Low-skilled manual	0.930*** (0.161)	0.841*** (0.165)		0.703*** (0.117)	0.612*** (0.118)		0.534*** (0.149)	0.353* (0.150)		0.212 (0.128)	0.316* (0.129)	
Unskilled manual	0.598*** (0.165)	0.516** (0.169)		0.449*** (0.116)	0.352** (0.117)		0.475** (0.152)	0.302* (0.154)		-0.238 (0.145)	-0.133 (0.146)	
Farmer	1.025*** (0.156)	0.885*** (0.165)		0.769*** (0.114)	0.618*** (0.118)		1.161*** (0.145)	0.832*** (0.151)		1.152*** (0.123)	1.100*** (0.127)	
SES (ref: Skilled manual)												
White collar		-	-0.980* 0.750*** (0.168)		-0.905*** (0.118)	-0.222 (0.308)		-0.174 (0.169)	-0.514 (0.456)		0.391** (0.125)	0.0715 (0.226)
Low-skilled manual		0.303* (0.148)	0.665 (0.363)		0.230* (0.102)	0.625* (0.254)		0.794*** (0.140)	0.723* (0.355)		-	-0.449* 0.501*** (0.113)
Unskilled manual		0.0931 (0.142)	0.0268 (0.351)		0.168 (0.102)	0.437 (0.259)		0.842*** (0.144)	0.863* (0.363)		-0.308* (0.125)	-0.229 (0.218)
Farmer		0.457** (0.157)	0.661 (0.388)		0.588*** (0.116)	1.297*** (0.289)		1.319*** (0.155)	1.717*** (0.390)		0.444** (0.146)	0.335 (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. OLS estimates of life expectancy 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*** (0.205)	0.538* (0.210)		1.186*** (0.148)	0.937*** (0.151)		0.522* (0.216)	0.181 (0.217)		0.556*** (0.162)	0.297 (0.163)	
Low-skilled manual	0.337* (0.164)	0.407* (0.165)		0.152 (0.120)	0.242* (0.121)		-0.452** (0.166)	-0.332* (0.166)		-0.304* (0.135)	-0.166 (0.136)	
Unskilled manual	0.0252 (0.168)	0.134 (0.169)		-0.183 (0.120)	-0.0747 (0.120)		-0.671*** (0.169)	-0.509** (0.169)		-0.758*** (0.154)	-0.597*** (0.155)	
Farmer	0.316* (0.159)	0.301 (0.162)		0.333** (0.119)	0.358** (0.120)		0.131 (0.162)	0.140 (0.164)		0.787*** (0.129)	0.845*** (0.130)	
SES (ref: Skilled manual)												
White collar		0.481** (0.167)	0.433 (0.411)		0.585*** (0.116)	0.285 (0.305)		1.418*** (0.163)	0.714 (0.487)		0.940*** (0.126)	0.572* (0.237)
Low-skilled manual		0.0423 (0.145)	-0.283 (0.353)		-0.157 (0.102)	-0.444 (0.261)		0.135 (0.142)	0.189 (0.375)		-0.439*** (0.121)	-0.0250 (0.219)
Unskilled manual		-0.545*** (0.153)	-0.608 (0.372)		-0.594*** (0.114)	-0.757* (0.295)		-0.288 (0.174)	-0.325 (0.481)		-0.462** (0.149)	0.153 (0.267)
Farmer		0.0775 (0.158)	0.124 (0.389)		-0.156 (0.120)	-0.195 (0.304)		0.462** (0.163)	0.0540 (0.437)		0.314* (0.154)	0.376 (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. OLS estimates of life expectancy 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*** (0.153)	-1.845*** (0.153)	-1.924*** (0.410)	-1.951*** (0.408)	-1.167*** (0.160)	-1.057*** (0.160)	-0.339 (0.460)	-0.292 (0.462)
1st quartile group	0.190 (0.115)	0.104 (0.115)	0.0410 (0.300)	-0.0254 (0.300)	-0.329* (0.130)	-0.284* (0.130)	-0.457 (0.376)	-0.438 (0.377)
3rd quartile group	-0.185 (0.116)	-0.0982 (0.117)	-0.266 (0.313)	-0.114 (0.315)	0.171 (0.128)	0.0585 (0.128)	0.367 (0.376)	0.335 (0.378)
4th quartile group	-0.659*** (0.133)	-0.319* (0.137)	-0.816* (0.373)	-0.434 (0.380)	0.916*** (0.140)	0.577*** (0.147)	0.313 (0.438)	0.226 (0.449)
SES (ref: Skilled manual)								
White collar		-0.0882 (0.149)		-0.546 (0.456)		1.160*** (0.149)		0.583 (0.495)
Low-skilled manual		0.745*** (0.126)		0.616 (0.355)		0.00772 (0.129)		0.205 (0.377)
Unskilled manual		0.818*** (0.130)		0.802* (0.365)		-0.369* (0.157)		-0.317 (0.484)
Farmer		1.548*** (0.138)		1.686*** (0.394)		0.581*** (0.147)		0.101 (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6. 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

Statistics Sweden (1915,1926,1935)

Table 7. Yearly consumption of alcohol, tobacco and selected foodstuffs.

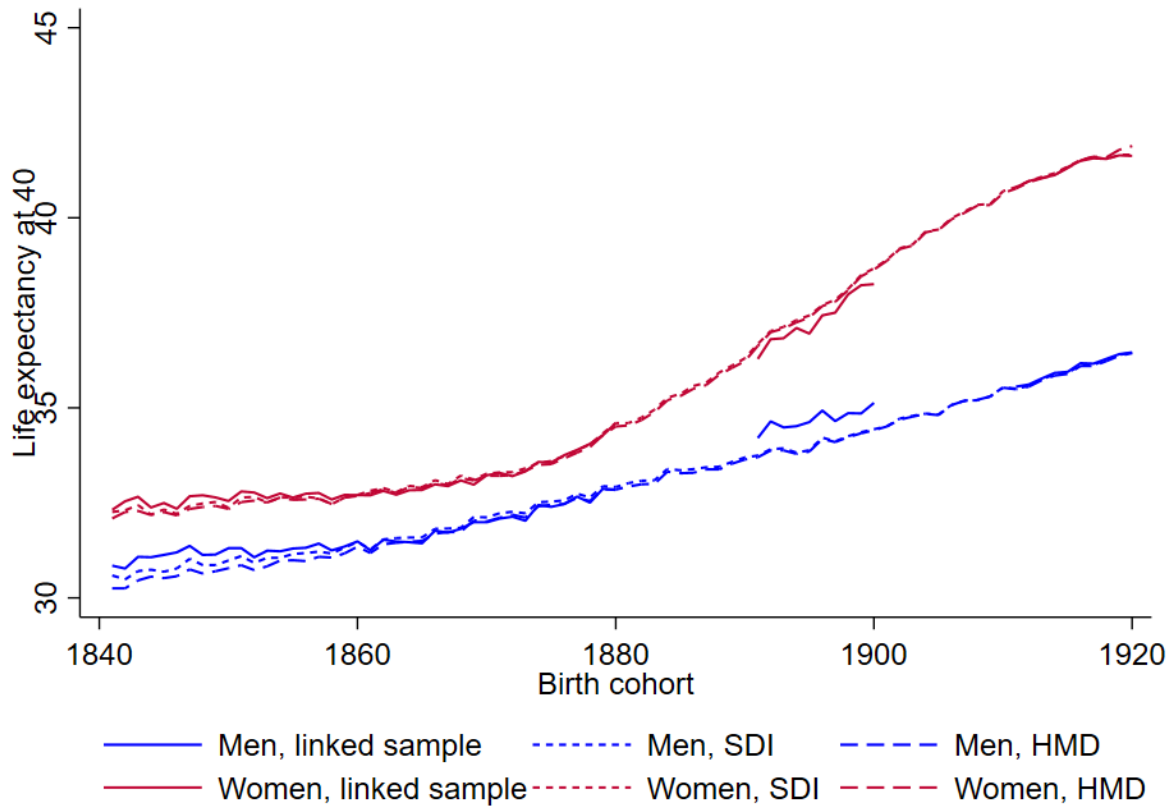
	Income per consumption unit (SEK)			
	<600	600-750	750-1050	>1050
<i>Panel A. Value (SEK)</i>				
Spirits	3.0	4.0	6.4	8.7
Wine	-	-	-	-
Beer	1.0	1.8	3.2	6.0
Tobacco	2.8	4.2	6.1	9.2
Cream	3.6	4.5	5.6	6.4
Wholemilk	26.8	35.9	40.6	49.9
Skimmed milk	7.2	5.6	4.8	2.0
Butter	20.7	26.1	31.3	36.7
Margarine	12.6	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
<i>Panel B. Share of total food, drink and tobacco expenditure (%)</i>				
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	-	-	-	-
<i>Panel C. Volume (litres or kilograms)</i>				
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

Appendix A: Additional Results

Figures

Figure A1. Cohort life expectancy 1841-1920 cohorts



Sources:

Tables

Table A1. Linkage rates by census year and age

Age	1880	1890	1900	1910	1930	1950	Total
<i>Panel A. Men</i>							
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
<i>Panel B. Women</i>							
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

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

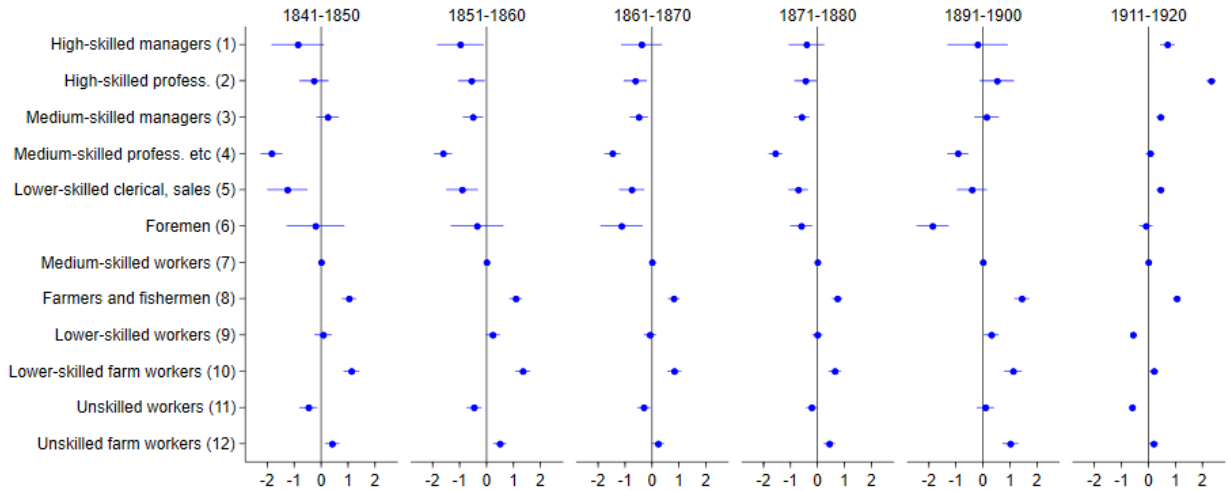
	1880		1890		1900		1910		1930		1950	
	Census	Linked	Census	Linked	Census	Linked	Census	Linked	Census	Linked	Census	Linked
<i>Panel A. Men</i>												
SES												
White collar	0.09	0.09	0.1	0.11	0.12	0.12	0.14	0.14	0.13	0.12	0.26	0.26
Skilled manual	0.09	0.1	0.11	0.11	0.13	0.13	0.16	0.16	0.14	0.14	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
<i>Panel B. Women</i>												
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

Appendix B: Sensitivity testing

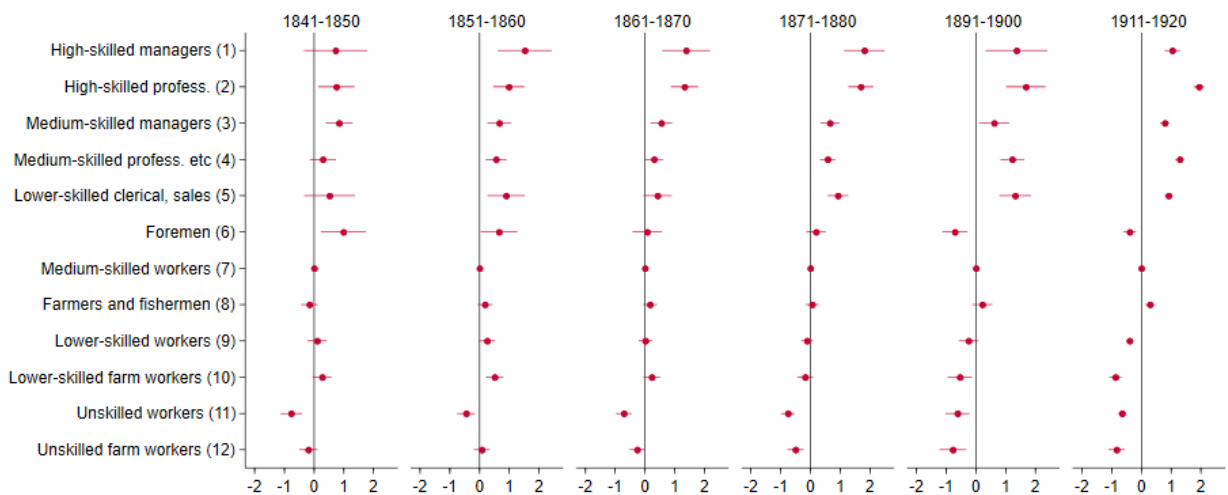
Figures

Figure B1. Coefficients from OLS estimates of life expectancy (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.

Figure B2. Coefficients from OLS estimates of life expectancy (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.

Tables

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

	1841-1850	1851-1860	1861-1870	1871-1870	1891-1900	1911-1920
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Men</i>						
<i>Dependent variable e50</i>						
SES (ref: Skilled manual)						
White collar	-0.728*** (0.129)	-1.073*** (0.113)	-0.914*** (0.098)	-0.799*** (0.084)	-0.023 (0.131)	0.835*** (0.048)
Low-skilled manual	0.452*** (0.114)	0.452*** (0.100)	0.243** (0.088)	0.338*** (0.074)	0.713*** (0.111)	-0.411*** (0.047)
Unskilled manual	0.094 (0.108)	0.085 (0.094)	0.094 (0.084)	0.248*** (0.074)	0.691*** (0.113)	-0.342*** (0.052)
Farmer	0.589*** (0.112)	0.683*** (0.102)	0.574*** (0.094)	0.607*** (0.083)	1.401*** (0.120)	0.875*** (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
<i>Dependent variable e60</i>						
SES (ref: Skilled manual)						
White collar	-0.523*** (0.111)	-0.782*** (0.096)	-0.663*** (0.085)	-0.447*** (0.072)	0.208 (0.115)	0.853*** (0.042)
Low-skilled manual	0.282** (0.097)	0.316*** (0.084)	0.151* (0.075)	0.226*** (0.063)	0.497*** (0.097)	-0.339*** (0.042)
Unskilled manual	0.122 (0.092)	0.116 (0.080)	0.190** (0.072)	0.171** (0.063)	0.529*** (0.099)	-0.305*** (0.046)
Farmer	0.392*** (0.096)	0.397*** (0.087)	0.348*** (0.080)	0.444*** (0.072)	1.092*** (0.105)	0.626*** (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
<i>Dependent variable e70</i>						
SES (ref: Skilled manual)						
White collar	-0.256** (0.095)	-0.375*** (0.083)	-0.104 (0.074)	0.039 (0.062)	0.429*** (0.100)	0.744*** (0.036)
Low-skilled manual	0.148 (0.083)	0.229** (0.072)	0.044 (0.064)	0.104 (0.054)	0.271** (0.084)	-0.232*** (0.035)
Unskilled manual	0.040 (0.078)	0.040 (0.068)	0.036 (0.062)	-0.003 (0.054)	0.212* (0.086)	-0.264*** (0.039)
Farmer	0.186* (0.082)	0.203** (0.074)	0.207** (0.069)	0.256*** (0.061)	0.670*** (0.091)	0.344*** (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
<i>Panel A. Women</i>						
<i>Dependent variable e50</i>						
SES (ref: Skilled manual)						
White collar	0.150 (0.138)	0.291* (0.116)	0.309** (0.101)	0.607*** (0.085)	1.334*** (0.128)	1.140*** (0.048)
Low-skilled manual	-0.064 (0.116)	0.192 (0.100)	0.129 (0.088)	-0.109 (0.075)	0.006 (0.113)	-0.357*** (0.049)
Unskilled manual	-0.428*** (0.117)	-0.197 (0.101)	-0.298** (0.092)	-0.504*** (0.084)	-0.307* (0.138)	-0.582*** (0.060)
Farmer	-0.367** (0.118)	-0.049 (0.106)	0.076 (0.097)	0.052 (0.089)	0.508*** (0.129)	0.354*** (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
<i>Dependent variable e60</i>						
SES (ref: Skilled manual)						
White collar	0.091 (0.118)	0.206* (0.098)	0.298*** (0.087)	0.611*** (0.073)	1.064*** (0.112)	1.053*** (0.043)
Low-skilled manual	-0.013 (0.098)	0.134 (0.085)	0.089 (0.075)	-0.126 (0.064)	-0.043 (0.099)	-0.284*** (0.043)
Unskilled manual	-0.229* (0.118)	-0.210* (0.106)	-0.181* (0.097)	-0.377*** (0.089)	-0.302* (0.129)	-0.525*** (0.070)

	(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
<i>Dependent variable e70</i>						
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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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

	1841-1850	1851-1860	1861-1870	1871-1870	1891-1900	1911-1920
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Men</i>						
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
<i>Panel B. Women</i>						
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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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)
<i>Panel A: Men</i>						
<i>P(survival past age 50)</i>						
SES (ref: Skilled manual)						
White collar	-0.004 (0.004)	-0.001 (0.003)	-0.003 (0.003)	-0.006** (0.002)	-0.002 (0.003)	-0.001 (0.001)
Low-skilled manual	0.005 (0.003)	0.008** (0.003)	0.001 (0.002)	-0.002 (0.002)	0.005* (0.002)	-0.001 (0.001)
Unskilled manual	0.000 (0.003)	0.000 (0.002)	-0.003 (0.002)	-0.004* (0.002)	0.006* (0.002)	-0.002 (0.001)
Farmer	0.010*** (0.003)	0.009*** (0.003)	0.006* (0.002)	0.004 (0.002)	0.008** (0.002)	0.004*** (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
<i>P(survival past age 60)</i>						
SES (ref: Skilled manual)						
White collar	-0.017*** (0.005)	-0.020*** (0.004)	-0.019*** (0.004)	-0.023*** (0.003)	-0.011** (0.004)	0.001 (0.001)
Low-skilled manual	0.014** (0.004)	0.015*** (0.004)	0.005 (0.003)	0.004 (0.003)	0.016*** (0.004)	-0.005*** (0.001)
Unskilled manual	0.000 (0.004)	-0.001 (0.004)	-0.005 (0.003)	0.001 (0.003)	0.015*** (0.004)	-0.004** (0.002)
Farmer	0.020*** (0.004)	0.022*** (0.004)	0.016*** (0.003)	0.012*** (0.003)	0.024*** (0.004)	0.015*** (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
<i>P(survival past age 70)</i>						
SES (ref: Skilled manual)						
White collar	-0.028*** (0.006)	-0.039*** (0.005)	-0.042*** (0.004)	-0.040*** (0.004)	-0.014* (0.006)	0.015*** (0.002)
Low-skilled manual	0.020*** (0.005)	0.020*** (0.004)	0.010** (0.004)	0.011*** (0.003)	0.028*** (0.005)	-0.012*** (0.002)
Unskilled manual	0.005 (0.005)	0.006 (0.004)	0.004 (0.004)	0.010** (0.003)	0.032*** (0.005)	-0.009*** (0.002)
Farmer	0.029*** (0.005)	0.031*** (0.005)	0.023*** (0.004)	0.023*** (0.004)	0.050*** (0.005)	0.031*** (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
<i>Panel A: Women</i>						
<i>P(survival past age 50)</i>						
SES (ref: Skilled manual)						
White collar	0.012*** (0.003)	0.013*** (0.003)	0.009*** (0.002)	0.007*** (0.002)	0.005* (0.002)	0.003*** (0.001)
Low-skilled manual	0.004 (0.003)	0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.000 (0.002)	-0.001 (0.001)
Unskilled manual	-0.003 (0.003)	-0.001 (0.002)	-0.007** (0.002)	-0.007*** (0.002)	-0.002 (0.003)	-0.001 (0.001)
Farmer	0.001 (0.003)	0.003 (0.003)	0.001 (0.002)	-0.001 (0.002)	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
<i>P(survival past age 60)</i>						
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
<i>P(survival past age 70)</i>						
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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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

	1841-1850	1851-1860	1861-1870	1871-1870	1891-1900	1911-1920
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: SES based on own occupation</i>						
SES (ref: Skilled manual)						
White-collar	-0.615 (0.527)	0.097 (0.394)	0.167 (0.290)	0.936*** (0.212)	2.297*** (0.244)	2.289*** (0.131)
Low-skilled manual	-0.583 (0.430)	-0.239 (0.344)	-0.059 (0.258)	0.094 (0.192)	0.579** (0.218)	0.620*** (0.129)
Unskilled manual	-3.237*** (0.517)	-1.895*** (0.426)	-2.228*** (0.328)	-1.769*** (0.254)	-0.315 (0.362)	-0.194 (0.158)
Farmer	-0.613 (0.678)	-1.127* (0.549)	-0.827 (0.505)	-0.000 (0.451)	0.138 (0.493)	0.483 (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
<i>Panel B: SES based on spouse's occupation</i>						
SES (ref: Skilled manual)						
White-collar	0.606*** (0.173)	0.749*** (0.147)	0.624*** (0.130)	0.746*** (0.112)	1.022*** (0.177)	1.045*** (0.055)
Low-skilled manual	0.013 (0.148)	0.200 (0.128)	-0.055 (0.115)	-0.332*** (0.100)	-0.375* (0.159)	-0.578*** (0.057)
Unskilled manual	-0.253 (0.145)	-0.070 (0.125)	-0.272* (0.114)	-0.491*** (0.106)	-0.578** (0.176)	-0.608*** (0.068)
Farmer	-0.268 (0.146)	0.052 (0.131)	0.106 (0.120)	-0.076 (0.110)	0.219 (0.166)	0.316*** (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B5. OLS estimates of life expectancy at age 40, 1861-1920 cohorts

	1841-1850		1851-1860		1861-1870		1871-1870		1891-1900	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Panel A: Men</i>										
SES (ref: Skilled manual)										
White collar	-0.740*** (0.150)	-0.572* (0.226)	-1.005*** (0.130)	-0.886*** (0.186)	-0.885*** (0.113)	-0.850*** (0.151)	-0.906*** (0.095)	-0.905*** (0.121)	-0.075 (0.148)	0.075 (0.172)
Low-skilled manual	0.587*** (0.132)	0.675** (0.208)	0.661*** (0.115)	0.631*** (0.170)	0.260* (0.102)	0.270* (0.138)	0.238** (0.085)	0.263* (0.108)	0.838*** (0.126)	0.771*** (0.144)
Unskilled manual	0.108 (0.125)	0.222 (0.196)	0.090 (0.109)	-0.047 (0.161)	-0.002 (0.098)	0.016 (0.133)	0.117 (0.085)	0.151 (0.109)	0.853*** (0.129)	0.778*** (0.148)
Farmer	0.868*** (0.130)	0.932*** (0.209)	0.898*** (0.118)	0.682*** (0.178)	0.711*** (0.108)	0.746*** (0.151)	0.680*** (0.096)	0.710*** (0.125)	1.592*** (0.137)	1.553*** (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
<i>Panel A: Women</i>										
SES (ref: Skilled manual)										
White collar	0.494** (0.162)	0.713** (0.274)	0.662*** (0.135)	0.649** (0.209)	0.561*** (0.116)	0.519** (0.158)	0.788*** (0.098)	0.804*** (0.123)	1.470*** (0.145)	1.389*** (0.165)
Low-skilled manual	0.056 (0.136)	-0.086 (0.238)	0.256* (0.117)	0.180 (0.185)	0.062 (0.102)	-0.046 (0.141)	-0.179* (0.087)	-0.214 (0.110)	-0.011 (0.128)	0.024 (0.143)
Unskilled manual	-0.490*** (0.138)	-0.653** (0.241)	-0.215 (0.119)	-0.289 (0.189)	-0.489*** (0.107)	-0.582*** (0.148)	-0.697*** (0.097)	-0.642*** (0.123)	-0.378* (0.156)	-0.288 (0.176)
Farmer	-0.301* (0.139)	-0.478* (0.243)	0.059 (0.124)	-0.108 (0.199)	0.106 (0.113)	-0.018 (0.158)	0.011 (0.103)	0.049 (0.132)	0.549*** (0.146)	0.557*** (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$