DISCUSSION PAPER SERIES

IZA DP No. 15064

Long-Run Mortality Effects of a Reform That Opened up Access to Secondary Education

Gerard J. van den Berg
Lena Janys
Kaare Christensen

FEBRUARY 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15064

Long-Run Mortality Effects of a Reform That Opened up Access to Secondary Education

Gerard J. van den Berg
University of Groningen, University Medical Center Groningen, IFAU Uppsala, IZA, ZEW and CEPR

Lena Janys
University of Bonn, HCM and IZA

Kaare Christensen
University of Southern Denmark and Danish Twin Registry

FEBRUARY 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

Long-Run Mortality Effects of a Reform That Opened up Access to Secondary Education*

We examine the effects of a major national schooling reform in Denmark in 1903, opening up access to secondary and higher education for poorer and for female children, on mortality, using individual-level records of Danish twins. We digitized education outcomes from historical registers and augmented these with data we digitized on parental socioeconomic status. The study design is combined with an exogenous indicator of economic conditions at birth to investigate whether education mitigates mortality effects of adverse conditions at birth. We find that the reform reduces mortality rates among males, notably those with a middle-class family background. Also, secondary education is less beneficial if conditions at birth are adverse. In general, the reform effect does not seem to be driven by improved information on healthy living but rather by a shift in social classes among the inflow into higher education.

JEL Classification: I1, I14, I20, N33, C41

Keywords: health, inequality, schooling, Correlated Frailty Model, twins

Corresponding author:
Lena Janys
Department of Economics
University of Bonn
Adenauerallee 24-42
53113 Bonn
Germany
E-mail: ljanys@uni-bonn.de

* We thank participants at an IZA Workshop on Health and Labor Markets, an American-European Health Study Group Meeting in Boston and a Workshop on Health and Inequality in Copenhagen, for comments and suggestions. Lena Janys was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy –EXC-2126/1–390838866 and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy –EXC-2047/1 –390685813. We thank Axel Skytthe of the University of Southern Denmark and the Danish Twin Registry for help with data merging, Kasper Kristensen and Tim Mensinger for excellent research assistance, and Niels Borger of the archive of the Danish Ministry of Education for providing access to the paper records of individual-level education data.

1 Introduction

The influence of education on non-labor market outcomes has been the subject of a vast body of work. This applies in particular to the impact of education on mortality and other health outcomes such as hospitalization and chronic diseases (see, for example, Cutler, Huang and Lleras-Muney, 2016). Much of this literature exploits relatively modest institutional changes in the education system, such as raising the minimum school leaving age. Such analyses cannot be straightforwardly generalized to more radical changes in access to education. In the context of the Covid-19 pandemic, with compulsory and prolonged lockdowns of schools, the question of how a complete reversal of educational access affects children's health outcomes, in the long run, has attracted renewed interest.[1]

This paper contributes to the literature by examining the effect of education on mortality in currently deceased cohorts exposed to a major national schooling reform in Denmark in 1903. This reform effectively opened up secondary and higher education access for poorer and female children. We find that education has a large and significant protective effect on men, reducing effective age by about 4.3 years, while there is no effect for women.

We use data that build on the administrative register of newborn Danish twins. Records are at the individual level and contain the days of birth and death, gender and zygosity of the twins, and information on individual early-life conditions that may confound the analysis if not taken into account. We augment these register data by digitizing key education variables from other registers and adding data on parental socioeconomic status derived from their occupation. Overall, the data enable us to study the effects of the reform and education on mortality, conditioning on gender, and other background variables. The occupation variables are used to shed light on social mobility effects. Also, using an exogenous contextual indicator for economic conditions at birth enables us to study whether education mitigates the mortality effects of adverse conditions at birth. Our identification strategy does not rely on the (relatively) strong assumptions necessary to interpret twin fixed-effects results as causal, although we use this as a robustness check. This study design has been criticized in the past (see ,e.g., Kaufman and Glymour 2011, Gilman, S.E. and Loucks, 2014) for potentially yielding biased estimates.

The provision of universal access to secondary education was one of the most extensive

[1]Indeed, as Grewenig et al. (2021) show that school closures, even with remote learning, indeed led to significantly fewer hours spent learning, resulting effectively in no schooling for students from a low-socioeconomic background.

policy reforms in developed countries in the 20th century. Plausibly, it led to significant shifts in society and social mobility. It is possible that this ultimately affected mortality rates as well. The results of this paper are thus of importance for any assessment of the benefits of such reforms and our understanding of changes in health outcomes over the 20th century. At the same time, they contribute to the literature mentioned above on education and health as well as to the literature on the long-run health effects of adverse conditions early in life (see, e.g., Almond and Currie, 2011, for an overview, and van den Berg, Doblhammer and Christensen, 2009, 2011, for applications with Danish twin data).

We collected and digitized the individual-level data on educational attainment from administrative paper records in the Danish Ministry of Education archives. Having individual-level information about life-span *and* educational attainment is essential because a large share of the discussion about trends in the gradient of education and mortality revolves around the question of whether these are due to changes in cohort composition, as, for example, pointed out in Currie and Schwandt (2016). Note that the key advantage of our merged register data is *not* that it contains twins. Instead, it is the fact that it contains individuals who were young around the time of the reform for whom we observe *both* the realized levels of education *and* the ultimate ages at death. Apart from our expanded twin data, no such data are available. The bonus advantage of the twin data is that it allows for fixed effect analyses of within-twin-pair variation in education.

Despite this relatively large amount of historical and life-course information at the individual level, the range of meaningful empirical analyses is restricted. First, Denmark around 1900 did not have a large population, so our data, even if they are population-wide, do not give rise to large samples. Per birth year, the number of twin members amounts to around 550, and of these, only a small fraction entered secondary education. Secondly, and more importantly, the 1903 reform of the educational system does not provide straightforward instrumental variation for inference on the effects of the individual education level. As we shall see, the roll-out was slow, and the length of time that students spent in higher education was not fixed. As baseline analyses, we, therefore, estimate models in which the exposure to the post-reform era is a covariate and separate models with indicators for whether the individual education level is high or not. In each case, we control for unobservables in a way that makes efficient use of the twin data by modeling the unobserved heterogeneity as a correlated frailty. This treats the effect of unobservables as a combination of an individual effect and a shared term between the twins. This, in turn, allows us to estimate the correlation between twins' unobservable characteristics separately for mono- and dizygotic twins, which is essential, as the former share more traits than the latter.[2] Following the usual line of reasoning in instrumental variable studies on the effects of education, we expect the effect of education to be bounded from below (in an absolute sense) by the coefficient for education in the twin survival model.

We use idiosyncratic macro-economic fluctuations to capture exogenous variation in economic conditions around birth. This is common in the literature on the long-run effects of early-life conditions (see, e.g., van den Berg, Lindeboom and Portrait, 2006) as these fluctuations are arguably exogenous drivers of household conditions. In our study, this approach is hampered by data limitations. The almost insurmountable task of manually tracking down twins in the historical education records forced us to restrict attention to a relatively small number of birth year cohorts around the cohort first exposed to the reform (1888–1897). The business cycle does not display significant fluctuations in this relatively short birth year interval (see also Figure 4.3). To proceed, we first consider a fixed effect approach in survival analysis, the so-called Stratified Partial Likelihood (SPL) estimation method (Ridder and Tunalı, 1999). This is the equivalent of fixed-effect panel data estimation in linear models. We assume that twin members share a fixed effect, and we include individual education and the interaction of individual education and the business cycle at birth in the specification of the mortality rate. The fixed effect effectively deals with the endogeneity of the individual educational choice, albeit at the assumption that it captures all systematic unobserved heterogeneity of both twin members, which has recently been called into question. Because of the latter, we also estimate survival models with correlated frailty terms within twin pairs and with individual education (or reform exposure) as well as the business cycle at birth and their interactions as covariates.[3] In that case, indicators of early-life conditions help to control for confounding of education.[4] Notice that an additional advantage of the SPL approach is that it deals with selection on unobservables due to changes in the composition of newborns across the cycle.

The approach in the previous paragraph improves on the existing literature that considers whether education can compensate for adverse early-life conditions. In that liter-

[2]Unobserved heterogeneity in the model deals with the fact that dynamic selection affects the composition of survivors as a function of age. If ignored, this tends to bias the covariate coefficient estimates towards zero (see, e.g., van den Berg, 2001, for an overview). This holds regardless of whether cohort-level or individual-level data is used and even if the variation in schooling is exogenous. We demonstrate the effects of ignoring unobserved heterogeneity and ignoring right- and left truncation the covariate effect estimates in Appendix A.2.

[3]Note that in the SPL approach, exposure to the post-reform education period is subsumed into the fixed effect as it is shared within twin pairs.

[4]Fletcher and Lehrer (2009) examine this by using a combination of genetic markers and sibling fixed effects as instruments for health problems in early childhood. They find that attention deficit disorders and early-childhood obesity adversely affect educational attainment.

ature, typically, at least one of the causal variables (education and early-life conditions) is not represented by an exogenous indicator, and fixed-effects approaches are not used to deal with this. For example, Fritze, Doblhammer and van den Berg (2014) interact exogenous business cycle fluctuations at birth with actual education in basic regression analyses of mental health later in life.

To understand our findings on the effects of the 1903 educational reform and to understand potential channels how education could affect mortality, we utilize a different part of the historical data archives on education. Rather than only using twins, we compare the social class distribution of the entire cohort in two years for two different secondary-school degrees: “Gymnasium” and “Realskole”.[5] The idea is that since these degree programs had different lengths, we can compare students who received their degree in the same year, where for one set of students (“Gymnasium”) it is a “before-after” reform comparison while for the other set of students (“Realskole”) it is an “after-after” comparison where any changes can be attributed to a general time-trend. Here, we find that social-class composition did change due to the reform, in line with a more prominent role of the middle class.

The remainder of the paper is structured as follows: In Section 2 we briefly discuss the existing literature with an emphasis on empirical evidence and general identification issues. In Section 3 we describe the institutional details of the educational reform that took place in Denmark in 1903. In Section 4 we describe the data used in this study. In particular, we describe the twin registry data in Section 4.1. In Section 4.2 we explain our method of collecting and matching the individual education data to the twin registry data on mortality in detail. In Section 5 we describe the empirical strategy, including the correlated gamma frailty model that is used to control for unobserved heterogeneity. In Subsection 6.1 we report all results of the estimations on the full sample and separately for male and female twins and discuss possible implications of these results, and in Subsection 6.2 we present the results on a changing social-class composition among degree-holders. Section 7 concludes.

[5] We digitized the entirety of the schooling records and extracted the information on the type of degree and the occupation of the father. Then we sorted these occupations into categories and used a social-class coding, a combination of rank and occupation.

2 Education and Mortality at Higher Ages

There is extensive empirical literature on the effects of education on health, with evidence from many countries. This body of work roughly falls into three categories. The first strand of literature attempts to link education and mortality. The second strand attempts to link education and other health outcomes such as hospitalization and chronic diseases. The third strand tries to uncover channels through which education may influence health outcomes in recent cohorts by examining health behaviors such as smoking and obesity. We briefly summarize the evidence, focusing on the first strand as it is most closely related to our study. Mortality is arguably the most objective measure of health, as it does not rely on self-reporting, which may be biased and have measurement error variances that are larger in certain subgroups. Measures of health other than mortality might suffer from confounding factors that bias coefficient estimates to zero. Hospital attendance or the number of physician visits could be positively related to education if highly educated people are more aware of potentially severe symptoms, which may still result in a lower mortality rate (if other medical conditions are diagnosed earlier). Higher hospitalization rates or more physician visits could only reflect a higher incidence of diagnosis instead of being an indicator of poor health. A recent study by Cawley and Choi (2018) finds that better-educated individuals report their health more accurately, even when reporting health behaviors such as smoking. Hernández-Quevedo, Jones and Rice (2005) find that when asked to report their health status, less educated people are less consistent in their answers.

Grossman (2016) provides an excellent overview of the schooling and health and mortality literature. He breaks down the evidence by health measure and whether the empirical method was suitable for estimating causal effects.

There are several potential channels through which education might influence health. The theoretical Grossman (2006) model predicts that information on the health hazards of lifestyle choices influences health behavior as an investment into good health. A higher education degree may also lead to higher income, enabling individuals to further invest in health goods (e.g., health insurance, healthy food). In addition, education may shift the occupational opportunity set faced by the worker. Occupations for low- and uneducated workers may be hazardous due to physical stress, handling of dangerous substances, and poor working conditions, which may be especially relevant for earlier cohorts. For our purposes, it is interesting to point out that this mechanism, if ignored, could mask the potential positive health effects of education. If workers have a comparative advantage in physical occupations due to their physical condition, there will be an adverse selection of workers into further education due to general health. This relates to the so-called healthy worker effect (for a review, see, for example, Ly and Sung, 1999).

The empirical evidence regarding mortality outcomes is not unequivocal, especially for men. Most work that allows for a causal interpretation of the results in this field examines cohorts born more recently than ours so that high-age mortality is often not observed. These findings are surprising, as educational gradients across different measures of health are well documented (Galama and Van Kippersluis, 2019). We hypothesize that many studies fail to find effects on mortality because of the issues mentioned above concerning unobserved heterogeneity and censoring. Also, twin-fixed effects studies, which are often used for causal identification, might be biased towards zero (Kaufman and Glymour 2011, Gilman, S.E. and Loucks, 2014).

Lleras-Muney (2005) exploits a compulsory schooling reform and changes in child labor laws in different states in the US. Interestingly, she finds that the IV estimates do not differ much from the OLS estimates. Albouy and Lequien (2009) exploit substantial increases in French compulsory education and find no effects. However, many spells in their data are right-censored. Madsen et al. (2010) adopt a twin study design, using a more recent cohort of Danish twins. Their findings indicate that the effect of schooling was strongest for the older cohorts of males. Behrman et al. (2010) also adopt a twin study design with the Danish twin registry data. They also face frequent right censoring due to a large fraction of the investigated cohorts still being alive at end-of-study. They fail to find significant effects based on twin differences. Malamud, Mitrut, Pop-Eleches (2021) examine evidence from an educational reform from Romania. They do not find any significant mortality effects of more years of education. However, the cohorts they examine were born in the late 1940s and early 1950s, i.e., they also face significant right-censoring problems, which potentially biases the effect toward zero. Halpern-Manners et al. (2020) also used a twin-fixed effects study design, but also for cohorts that were much more recent. They do find a modest effect of education on age-at-death but are potentially exposed to biased estimates due to the fixed-effects design and the underlying assumptions, as well as concerns about right-censoring.

3 The Reform

We exploit an exogenous source of variation in the form of an extensive schooling reform that was introduced in 1903 in a law called the "Law for higher comprehensive schools" ("Lov om højere almeneskoler").

Previously, the Danish schooling system was based on compulsory six-year education, starting in the calendar year a student turned seven. The elementary school providing this education was free and ended without a degree. After this, there were two options to obtain a degree in two schooling tracks. The "Realeksamen" was obtained at the "Realskole" at approximately age 15/16, while the Gymnasium led to the "Studentereksamen" at about age 17/18, which was a prerequisite to study at university. Crucially, the gap between the end of the elementary school at age 12 and entering either the Realskole or Gymnasium had to be bridged with a private education which only wealthy families could afford. Additionally, both the Gymnasium and the Realskole had entrance requirements not taught in public schools, such as Latin and Greek.

The Ministry of Education (Undervisningsministeriet) described it: "Before 1903, there was no coherent connection between elementary school and further education"(Schmidt, 2005). The purpose of the reform was to provide students with a bridge between elementary school and further education. The reform law was passed in 1903. Figure 1 displays the most important elements of the reform. See Skovgaard-Petersen (1967) and Schmidt (2005) for detailed descriptions (in Danish) of the reform and its enactment.

The reform obligated elementary schools to extend grade levels beyond the sixth grade (i.e., beyond age 12) and introduce a free middle school. They were allowed to do so step by step, introducing 7th grade first and then continuing upwards, adding a grade level each year. Therefore, the first year where the Studentereksamen was administered under the new regime was in 1910 (specified in the law). Since students had to be at least 17 years old to be allowed to take this exam, it follows that the first cohort affected by the reform was born in 1893.

There is reason to suspect that the reform was not carried out equally across the country. In their report, the Ministry of Education noted that "it was especially city/town children who had the opportunity to follow the way through the middle school and further to other qualifications because middle schools were almost exclusively established in cities and towns." (see Schmidt, 2005).

It is important to note that the post-reform middle school track did not automatically

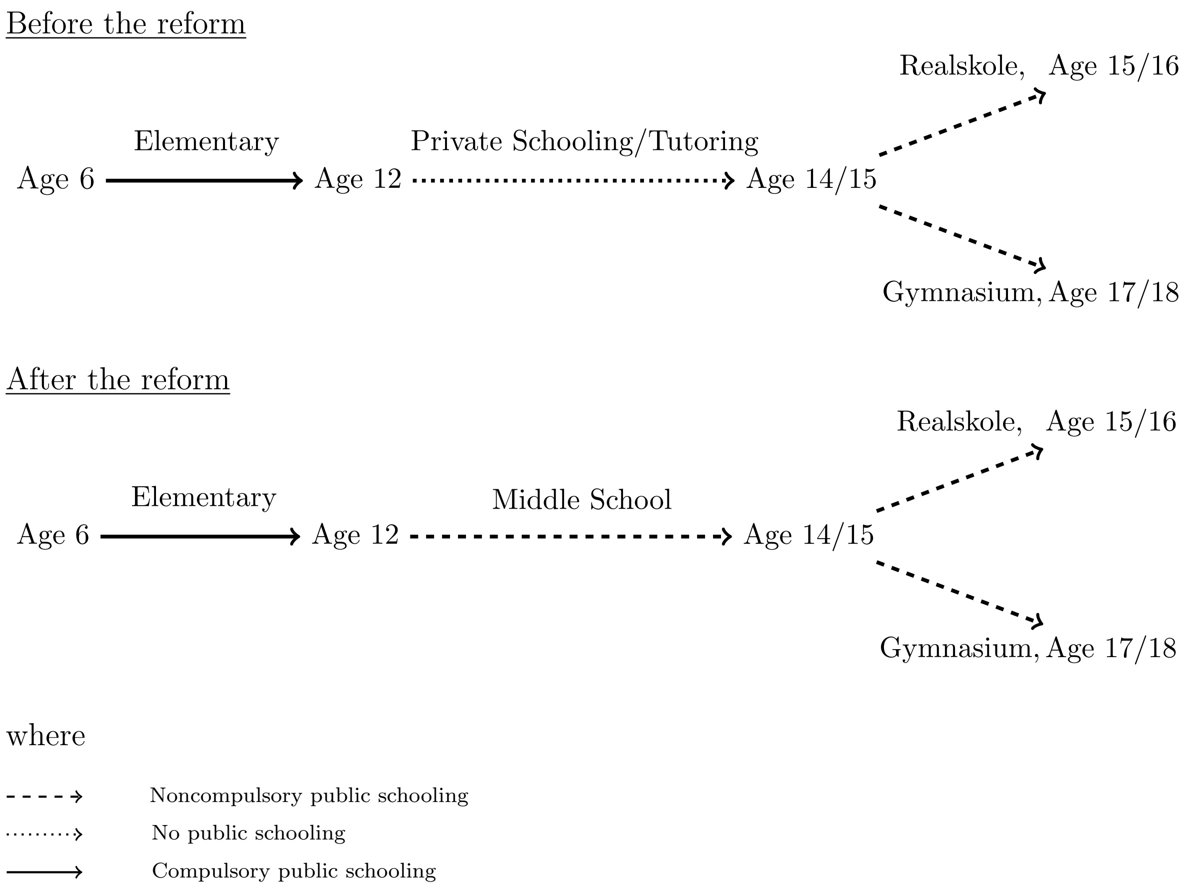

Figure 1: **Visualization of the basic elements of the 1903 educational reform.**

lead to completion until and including a degree (which is our education outcome variable). An alternative scenario is that students acquired additional years of education after finishing compulsory schooling and then dropping out without an exam. There is evidence of such attrition: just between the grades three and four of the middle school, the number of enrolled students in grade three in 1912 dropped from 5401 to 4500 in grade four in 1913 (see Danmarks Statistik, 1914). Consequently, regressing the probability of obtaining a degree on being exposed to the new system after the reform will underestimate the reform's average effect on educational attainment.

The reform introduced substantial improvements in the formal education system for girls. The law explicitly includes that middle school should be equally accessible for boys and girls. Furthermore, the reform gave girls the same right as boys to gain entrance to the Gymnasium, and thus this reform effectively enabled girls access to higher education. Even though girls had been granted the right to attend university in 1875, Girls had not been allowed access to the Gymnasium, and there was only one Girls' school where girls could take the "Studentereksamen". The reform also established a Girls school exam, comparable to the "Realeksamen" with fewer maths and more language requirements. Previously, girls had often left girls' schools without any degree. The uptake was high. We present some descriptive statistics for the aggregate sample of all students. For example, the number of girls taking the "Studentereksamen" rose by 100 percent from 1909 to 1910 (see Table 6 for aggregate statistics of the entire population on passed Studentereksamen), see also the Danmarks Statistik (1914) publication on degree-awarding schools from 1912 for details.

A limitation of using the reform for inference originates from its sluggish implementation. There was no immediate dramatic increase in the number of pupils who received a degree. Those students most likely to be exposed to the new regime in the first years after the reform lived in cities and towns and not in rural areas. One may suspect that the complying students had relatively low opportunity costs of staying in school. Academic entry requirements for the middle school further restricted access. In relative terms, the national increase in schooling uptake was significant (see Table 6), but the absolute number of pupils who graduated with a degree was small and remained so after the reform. This means, in particular, that in our twin sample, the surplus enrollment after the reform can be expected to be low, at least in the post-reform years in our observation window. This, in turn, restricts the scope for IV analyses based on the reform. We report the results from the first stage of regressing the exam on the reform, with and without additional covariates in Table 7.[6]

4 Data

4.1 Twin Data

The Danish twin registry contains detailed information on mortality (including the exact day of birth and death and the cause of death) and information on some initial conditions at birth which includes the general area (Jutland, Funen, Sealand, and Copenhagen), the degree of urbanization (rural, town or Copenhagen) and the zygosity (i.e., identical twins (monozygotic) or not (dizygotic)) of the twins. For a detailed overview of the entire data set and data collection methods, see Skytthe et al. (2002). We restrict our sample to twins born four years before and four years after the cut-off for exposure to the reform.

[6] Despite the reform being more salient in towns and cities, the key findings of the first-stage regression do depend on the degree of urbanization in the place of birth.

Therefore, we restrict the sample to twins born in 1888-1897. The first birth cohort exposed to the reform was born in 1893. Overall this comprises 5,695 individuals who survived until 1943 (the first year that reliable mortality records are available). If we only keep observations with status information on both twins, then we have mortality records for 2559 twin pairs. This sample size, though modest, is still significantly larger than that of comparable studies, e.g., Halpern-Manners et al. (2020). A survey from 1966 added more detailed information on educational status, family background, occupational history, and other socio-economic information to the twin registry data. However, relying on information from this survey would entail restricting our sample to individuals who lived at least until 1966 and who responded to the survey. This would severely limit the number of observations since the first cohort we want to consider in our analysis was born in 1888. In Section 4.2 we describe in detail how we instead collected the individual-level educational data and how we merged them with the data from the Danish twin registry.

Table 2 shows the proportion of twins for whom we have reliable status information, that is, for whom we observe the exact time of death or the year of emigration. There are only a few missings. Most efforts to uncover complete information were made for monozygotic twins, where the share of twins recorded to have died is 95.63%, with 3.81% emigrated. The second priority in uncovering mortality information was same-sex twins in general. The lowest effort was made for different-sex twins, where accordingly, the share of twins with an unknown status is the largest (36.37%). However, there are still ways to utilize the information that we have on these twins: we can censor most of the missing observations since, in 1943, the Danish Mortality Register started recording deaths systematically. All deaths that occurred before that were not subject to systematic recording. More efforts were made to uncover the time of death even before that date for monozygotic twins. We censor observations with the exit status emigrated at the date of emigration[7]. Tables 4 and 3 show some descriptive information about the twin data set.

4.2 Education Data

Because the relevant birth cohorts used in this study are comparatively old, the survey information on the educational status is scarce. However, the archive of the Danish Ministry of Education (Formerly the Danish Ministry of Education and the Church) contains paper records of all students who completed a schooling degree for each year from 1848 until the records were digitized. These schooling records, called “Meddelelser

[7]We only know the year of emigration, so the date of emigration was set as July 1st in that year.

om de hoejre Almeneskoler", are comprehensive and contain information on the student's complete name, date of birth, place of birth, their final score, and the subjects that they graduated in. We digitized these paper records by scanning the pages individually and then extracted the names and birthdays with a text recognition program, which can recognize textual information from a scanned document.

Appendix B shows one scanned example page and the same page after it was treated with the text recognition program. After extracting the data, we merged these digitized names with the information from the twin registry. We used a database program and pre-sorted on last name (including common spelling variations such as Christensen and Kristensen) and birth year and then matched twins that we found by hand (by comparing the exact day of birth and, in uncertain cases, the place of birth by hand). This laborious procedure ensures that the twins we did not match did not receive a schooling degree.

4.3 Additional data on economic conditions

We supplement our main data set with data on economic conditions around birth, which have been shown to impact health (see,e.g., Berg, Doblhammer and Christensen, 2009, 2011). We show how this indicator varies over the relevant time horizon considered in our study in Figure 4.3.

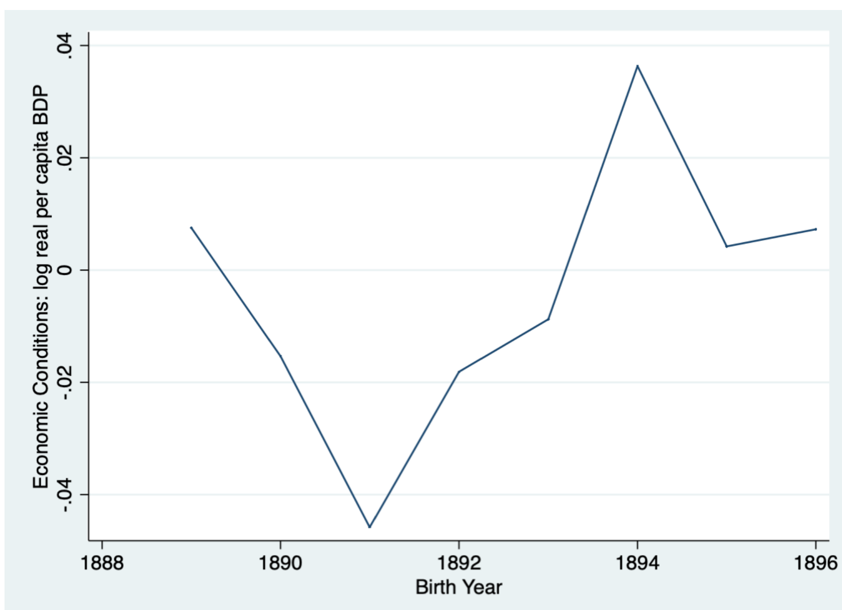

Figure 2: Economic conditions over all birth years.

5 The Empirical Model

This section introduces the empirical model and explains how we control for unobserved heterogeneity, which is crucial when mortality is the outcome of interest, as dynamic selection causes a bias towards zero, even if education were completely exogenous. The derivation of the likelihood function and complete parametric specification of the functional forms are in Appendix A. Since we have the exact date of birth and the exact time of death, we have continuous information on lifetime duration so that we treat *lifetime* as a continuous random variable. Using a duration model instead of OLS estimates or binary estimators has the advantage of facilitating dealing with censoring, e.g., due to migration or loss of follow-up and with left-truncation, which could lead to biased estimates due to dynamic selection. It allows for a straightforward interpretation of the covariate coefficients. In Appendix (A.2), we illustrate the problem of dynamic selection by way of a simulated example.

As a starting point, we model the individual mortality (hazard) rate as a simple mixed proportional hazard model. The individual hazard rate is a function of individual characteristics at birth x and a dummy variable that indicates a schooling degree S.

$$\log \theta(t|S, x) = h(t) + \beta' x + \eta' S \tag{1}$$

The parameter of interest, therefore, is η. The function $h(t)$ is called the baseline hazard. We assume a conventional Gompertz functional form for the baseline hazard, with parameter B,

$$h(t) = B \cdot t \tag{2}$$

Alternatively, we may estimate an "intention to treat" effect of the reform on the entire population of twins. As described earlier, the effect of the reform was widespread because more people had access to schooling beyond sixth grade but not necessarily obtained a subsequent degree. Therefore, the reform may have affected unrecorded further education on mortality. Let R be a dummy variable indicating whether or not a person was born in a year that exposed them to the reform, that is, born in or after 1893. Equation (3) is the reduced form equation,

$$\log \theta(t|R, x) = h(t) + \beta' x + \omega' R \tag{3}$$

The covariates x are *region of birth* (Jutland, Funen, Sealand excluding Copenhagen, and Copenhagen), *degree of urbanization* (Copenhagen, Town and Rural), a term $\log(birthyear - 1887)$ and sex and zygosity. The term $\log(birthyear - 1887)$ captures smooth effects of slow secular changes in society. Note that including age and a log-linear function of the birth year in the mortality rate entails that these terms also jointly capture smooth long-run effects of current calendar time on mortality. We also include a business cycle indicator for the year of birth in our specifications. This indicator is derived from a decomposition of the time series of log real per capita GDP and is taken directly from van den Berg, Doblhammer and Christensen (2011).

Unobserved Heterogeneity. As alluded to before, not controlling for unobserved heterogeneity leads to potentially biased estimates in duration models due to dynamic selection. In In Appendix (A.2) we discuss the reason for this in detail. Frailty terms V capture the effects of unobserved covariates on the individual hazard rate in the context of a proportional hazard model. The frailty has a multiplicative effect on the individual hazard rate,[8]

$$\log \theta(t|S, x, V) = h(t) + \beta' x + \eta' S + \log V \tag{4}$$

It is plausible that V is correlated within twin pairs. We assume it to be the sum of a shared part W and an individual part V_i^0,

$$V_i = W + V_i^0 \text{ where } i = 1, 2. \tag{5}$$

Following, e.g., Wienke (2003), we take the joint distribution of V_1, V_2 to have a Cherian bivariate gamma distribution, where the shape and scale parameters of the gamma distribution are equal. The marginal distributions are identical, and the mean is normalized (i.e., included in the constant term of $\beta' x$). Accordingly, the joint distribution has two parameters: the variance σ of V_i and the correlation ρ of V_1, V_2. The correlation can be expressed as the fraction of the variance of V_1, V_2 that is explained by the variance of the shared term W:

$$\rho = \frac{var(W)}{var(V_i)} \tag{6}$$

Estimation of the correlated frailty model includes the estimation of ρ and σ.

[8] For an introduction to frailty modeling, see, e.g., Wienke (2003).

The bivariate duration density function is given by:

$$
\begin{aligned}
f(t_1, t_2) \quad = \quad & \frac{(1-\rho)^2 \bar{F}(t_1)^{-\rho} \bar{F}(t_2)^{-\rho} f(t_1) f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_1)^{-\sigma^2} - 1)^{\rho/\sigma^2}} + \frac{(1-\rho)\rho \bar{F}(t_1)^{-\rho} \bar{F}(t_2)^{-\rho-\sigma^2} f(t_1) f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+1}} \\
& + \frac{(1-\rho)\rho \bar{F}(t_1)^{-\rho-\sigma^2} \bar{F}(t_2)^{-\rho} f(t_1) f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+1}} + \frac{\rho(\rho+\sigma^2) \bar{F}(t_1)^{-\rho-\sigma^2} \bar{F}(t_2)^{-\rho-\sigma^2} f(t_1) f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+2}}
\end{aligned}
\tag{7}
$$

where $\bar{F}(t_i)$ is the survival function of the univariate frailty model and $f(t_i)$ is the univariate density function of the univariate frailty model. The derivation of the bivariate likelihood function for the correlated frailty model is in, e.g., Yashin, Vaupel and Iachine (1995). Appendix A lists the parts of the likelihood function that we use in the estimation, including the likelihood contributions for censored and partially censored pairs. Our data also suffers from left truncation. This is because we only observe individuals who survived until at least 1943, the first year that reliable mortality records were kept. As we document in our simulation example in Appendix (A.2), both truncation and right censoring can potentially bias coefficient estimates.

One may consider augmenting the above model with a first-stage IV equation linking the reform to the realized level of education. As we already alluded to in Section 3 and earlier in this section, the data are insufficiently informative to pursue this. Hence, as mentioned in the introduction, we treat the reform as an intent-to-treat variable. Moreover, following the usual line of reasoning in IV analyses on effects of education, we take the causal effect of education to be bounded from below (in an absolute sense) by the estimated coefficient for education in the above mortality rate model.

6 Results and Discussion

6.1 Results on Education and Mortality

Table 1 shows the main estimation results for both $exam$ and $reform$. Full results can be found in Tables 8 and 9.

The full table with all controls 8 reports the parameter estimates for the correlated frailty model with the realized level of education included in the vector of covariates x in column (1). The parameter ρ_{dz} is the correlation between unobserved heterogeneity within dizygotic twin pairs; ρ_{mz} within monozygotic pairs. Columns (2) and (3) show the results for males and females, respectively.

The exam coefficient is negative, implying that having a schooling degree decreases mortality risk and is significantly different from zero. The results by gender reveal that this is entirely due to the effect among males. A degree significantly reduces their mortality risk; the coefficient is large and precisely estimated. For females, the coefficient is negative, but it is insignificant. These results are mirrored when eligibility for the reform is used instead of having a schooling degree. The results are reported in Table 9, column (1) again reports the results for the whole sample, columns (2) and (3) for men and women separately. The overall coefficient is also negative (meaning that exposure to the reform causes the mortality risk to decrease), but not quite as large as education, which is in line with the "intention to treat" interpretation. The coefficient is not as large for males, but it is still sizeable and significant. For females, the estimated reform effect is positive, but again, it is entirely insignificant, even though the reform explicitly targets girls' educational opportunities.

In terms of magnitude, the effects for male twins are quite large. As a back-of-the-envelope calculation, we follow Kulinskaya et. al (2020) and calculate an approximate reduction in effective age using $\Delta t_j \approx \log(\mathrm{HR}_{exam}) / \log(B)$. Having an exam decreases effective age by approx. 4.32 years averaged overall characteristics.[9] We could also interpret the coefficient value as a probability (on a small interval), and from the life-table, calculate the baseline hazard of death at particular ages: At age 73, the baseline probability of death is 0.044. Having an exam in this context means a reduction of that probability to 0.028, which is a sizeable difference.

The correlation of the variances of the unobserved heterogeneity term ρ_{mz} in the full sample is estimated to be 0.955, which suggests that for monozygotic twins, the use of a shared-frailty model is justified, while ρ_{dz} is 0.827, which supports the use of the correlated-frailty model instead of the shared-frailty model. To interpret the estimate of B, note that we rescaled age in days by dividing it by 10,000.

Cutler, Huang and Lleras-Muney (2015) analyze data covering centuries with multiple educational reforms and economic fluctuations and find that education has a protective effect for those graduating in bad times. We, therefore, include an interaction between education and a recession indicator for the graduation year, but the corresponding coefficient is not significant.

Next, we consider the role of economic conditions around birth. The business cycle indicator at birth represents these. Table 8 does not suggest any effects. To proceed, we estimate a fixed-effect duration model, exploiting within-twin variation using the stratified

[9] $(log(0.683)/(2.41106) * 10000/365)$

	(1) Whole Sample	(2) Females	(3) Males
Exam	**−0.38****	**−0.31**	**−0.50****
	(0.17)	(0.23)	(0.25)
Reform	**–0.10**	**0.11**	**−0.28***
	(0.11)	(0.16)	(0.15)
Controls	Yes	Yes	Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, for selected parameter estimates.

Table 1: *Regression main results*: Correlated frailty model with exam and reform exposure as an explanatory variable. Column (1) reports the estimates for the whole sample, columns (2) and (3) show the results for males and females, respectively. Full results can be found in Table 8 and Table 9.

partial likelihood estimation method. This assumes that the unobserved heterogeneity term is identical within twin pairs but relaxes the assumption that unobserved confounders are independent of the included observed covariates. Hence, this approach deals with the potential endogeneity of the level of education by allowing it to depend on an unobservable twin-pair-specific fixed effect. The determinants of mortality that do not vary within twin pairs cancel out of the estimation, and therefore their effects cannot be estimated. This also applies to the early-life conditions themselves, as they are identical within twin pairs. However, the marginal effect of education and the interaction effect between conditions at birth and education are identified. This constitutes an innovative approach to study interaction effects of early-life conditions and education on mortality.[10]

The estimation results in Table 10 confirm that secondary education is protective against mortality. However, the protective effect is absent if economic conditions at birth are adverse. Thus, secondary education does not mitigate adverse conditions at birth.

The results of this fixed-effects exercise are based on a sample of same-sex twins born in 1888-1997 (the sample size equals 2012 individuals). The birth years are relatively homogeneously distributed across 1888-1897. Both here and in general, in this subsection,

[10]This fixed-effect approach cannot be used to study the reduced-form or ITT effect of the reform as the latter has identical values within twin pairs. Also, we do not use it for the business cycle in the year of graduation as the latter is less predetermined than the year of birth, and hence the level of education cannot be manipulated to mitigate its effect.

it holds that if we include different-sex twin pairs or if we modify the birth year window, then the findings do not substantially change.

We now return to the differences in the results by gender. A possible explanation for the apparent absence of an effect among females is that the subsample of girls taking an exam is too small. We try to evaluate this by simulating a more extensive data set, using the estimated model while increasing the number of observations by expanding the sample size. Even if we quadruple the number of observations, we do not find any significant effect of having a schooling degree on the mortality rate for females.

To proceed, we may interpret the difference between males and females as informative on the channels through which education delays mortality. There are not many channels where we should expect large gender differences. The "information channel" (individuals making better-informed choices about their health) should work in the same direction. At least there seems to be no reason why information effects would be much more pronounced for males. Another possibility is that, in that era, education enabled men to make different occupational choices, shifting from physically demanding work to professional work, which can have large effects on life expectancy.

Women's labor force participation at the onset of the 20th century was low, and the single highest mortality risk factor for women of childbearing ages was giving birth. Even in Sweden, which was comparable to Denmark in having some of the lowest maternal death rates in the world, there were still about 250 maternal deaths per 100.000 live births at the onset of the 20th century (see Loudon, 2000). A surprising finding of Loudon (2000) is that the risk of maternal death was inverse to social class, meaning that obtaining a higher social class through education might have increased the risk of dying in childbirth. Van den Berg, Gupta, and Portrait (2010) examine fertility and early life conditions as factors for mortality in cohorts of Dutch women born towards the end of the 19th century and find that education delayed the time of first fertility in these earlier cohorts and that age at first birth was positively related to the mortality hazard. They also show that maternal mortality was the leading cause of death and responsible for up to 11% of all deaths among women aged 20-49.[11] It is reasonable to assume that, while death is an extreme outcome, there are other complications (such as fistulas and infections) that can increase mortality for some time after the latest birth.

[11]In contrast, today, birth complications are not even among the ten leading causes of death among women in that age bracket, according to Statistics Denmark.

6.2 Social Class

This subsection aims to shed some more light on how the reform affects mortality outcomes for males. In a "first-stage" estimation, i.e., regressing the probability of having a schooling degree on being eligible for the reform, the reform coefficient is not statistically different from zero once a time trend is added. However, we know from Table 6 that, at the national level, enrollment did respond to the reform, so the lack of a "first-stage" effect in the twin data must be due to the small sample size.

Also, recall that in the twin data, the reform *does* similarly affect male mortality as it affects individual schooling, albeit with a smaller coefficient, even when smoothly controlling for time-trends. To interpret this, we should keep in mind that some educational effects of the reform may not be visible in our education variable. First, recall evidence that the reform managed to keep students in school longer without obtaining more degrees. Since the reform was not targeted at degree schools, the latter schools might not have kept up with an increase in eligible students. Therefore it is possible that the average student went to school for a more extended period, something we cannot observe in our data.

Second, higher education might have shifted the distribution of social class among the students who continued their education to obtain a degree. To assess this, we analyze a specific subsample of our original education data using the entire cohort of students in Denmark instead of only twins. The schooling information described in Section 4.2 also reports the father's occupation at the time of graduation, albeit we only have this information for students who received a degree. We sort these occupations into nine social class categories. In the absence of a reliable Danish classification, we use the Swedish "Klassificeringschema för socio-ekonomiska grupper" that is based on occupation codes (*Yrkeskod*) and occupational classes (self-employed, employee, worker). The nine social classes roughly contain the following. (1) and (2) are agriculturally related occupations with (1) for landowners and (2) for workers. (3) are self-employed workers. (4) are self-employed variants of code (6). (5) are administrative workers for the government. (6) are priests, lawyers, academics, doctors, police, and similar professionals, not self-employed. (7) are manual workers. (8) are service workers. Finally, (9) are the military. In Appendix B.2 we describe the categories in more detail.

To separate general time trends from the reform's effect, we compare two different groups in two different years. The first group consists of students who received a "Realskole" degree, i.e., a short degree obtained one year after middle school. These cohorts of students graduated under the new regime in the years we consider, 1909 and 1910. The second group consists of the students who received a "Studentereksamen"; these students were exposed to the old regime if they did the exam in 1909 but to the new regime if they did it in 1910. If the social class composition had changed due to the reform, the social class composition for the "Gymnasium" category should change, whereas the social class composition for the "Realskole" category should only be affected by the general time trend. Figure 6.2 shows the results, with the upper panel representing the relative changes and the lower panel showing the absolute frequencies. The lighter shaded bars represent the "Gymnasium" category while the bars in blue indicate 1909. The x-axis depicts the different social class categories. Comparing the relative composition of the social classes in the upper panel of the "Realskole" for both years (the solid colored bars), we see that there are indeed some shifts in the relative composition of the students. Notably, there are opposing trends for Gymnasium and Realskole in social class (3), which comprises smaller entrepreneurs (such as grocers). While there is an increase in social class (3) among the Gymnasium graduates, the opposite is true for Realskole graduates, which suggests a shift in social class composition due to a rising middle class. The same can be argued for social class (4) ("self-employed professionals") and (5) ("employed executives").

7 Conclusion

This paper estimates the effect of far-reaching education reform in 1903 on mortality, using individual-level data on both education level and mortality.

We find that the reform and the level of education have a protective effect on mortality among males but not among females. Among the latter, in that sense, the reform did not lead to subsequent changes in the life course. This is a sobering conclusion in the light of the stated intentions of the reform, which included opening up a path to higher education for women. Of course, Denmark around 1900 differs from modern societies with advanced economies and high mobility. At the time, the life course of females may be predetermined to a more considerable extent by social background, and it may be that education could only play a limited role. In that sense, society may be more similar to that in a developing country with large gender inequality. Apparently, creating access to higher education in such a society is not sufficient for most females to enhance their health at higher ages, at least when this is measured by later-life mortality. We make two (related) caveats here. First, access to higher education may simply take decades or more of calendar time before

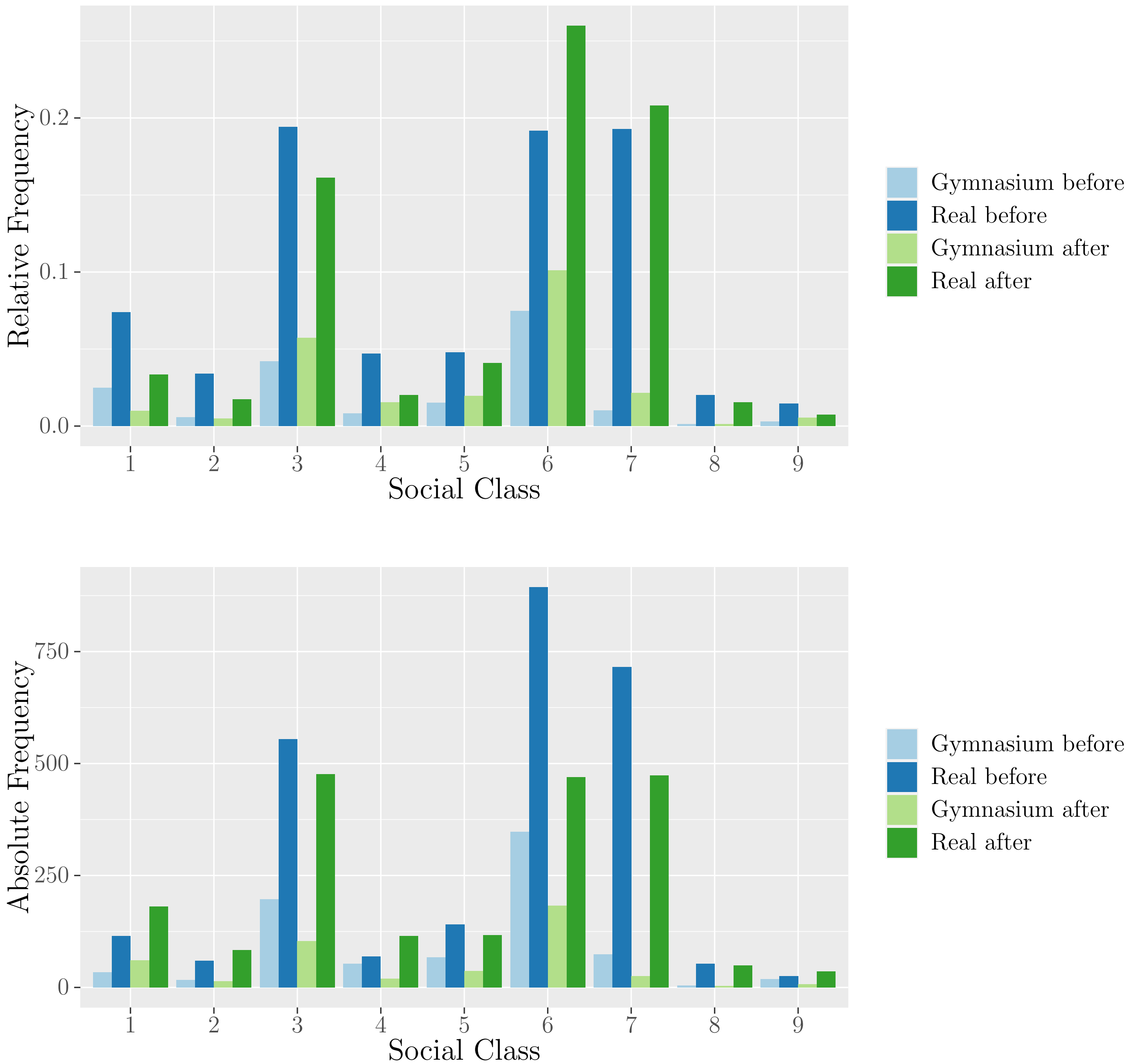

Figure 3: The social class composition of the different schooling degrees for before and after the reform (for "Gymnasium"), for "Realskole" both years correspond to eligible cohorts. The nine social classes: 1,2: Agriculturally related occupations with 1: Landowner, 2: Worker. 3: Self-employed workers, 4: Self-employed variant of code 6; 5: Administrative work for the government; 6: Priests, Lawyers, Academics, Doctors, Police, etc., not in self-employment, 7: Manual workers, 8: Service workers, 9: Military. **Upper Panel**: Relative frequencies and **Bottom Panel**: Absolute frequencies

female cohorts benefit from them across the board. Secondly, access to higher education may not be a sufficient condition, but it may be a necessary condition that needs to be augmented by other reforms or followed by demographic changes in fertility patterns. Our data cannot address these caveats but provide interesting avenues for further research.

Among men, we find evidence that those from a lower (but not the lowest) socio-economic background move more often to higher education (often just for some additional years, without necessarily completing the final exam and without continuing into university). This makes their subsequent occupations substantially different from those of their fathers, being less industrial and more oriented towards better-paying jobs in the service sector. This is consistent with a reduction in their mortality rate later in life. Following this line of reasoning, the reform seems to have served its purpose for young males from modest social backgrounds. It enabled them to follow higher education, leading to lower mortality rates.

As a by-product of the analysis, we find that secondary education is less beneficial for mortality reductions if economic conditions at birth are adverse and more beneficial if those conditions are favorable. Thus, secondary education does not mitigate adverse conditions at birth. In this sense, reforms that open up access to secondary education may have an intergenerational multiplier effect for those who manage to enroll.

Although the results are based on a historical reform, they may be used to predict the implications of prolonged reductions in schooling, as was the case in much of the world during the recent pandemic. Our results suggest that children from a lower socio-economic background may not fully recover in terms of their attained level of education and hence may suffer from higher mortality rates many decades after the lockdown.

References

Albouy, V., Lequien, L. (2009), Does compulsory education lower mortality?, *Journal of Health Economics* 28(1), 155–168.

Almond, D. and Currie, J. (2011), Killing me softly: The fetal origins hypothesis, *Journal of Economic Perspectives* 25(3), 153–172.

Behrman, J.R., Kohler H., Jensen, V.M., Pedersen, D., Petersen, I., Bingley, P., Christensen, K. (2010), Does More Schooling Reduce Hospitalization and Delay Mortality? New evidence based on Danish twins., *Demography* 48(4), 1347–1375.

van den Berg, G.J. (2001), Duration models: specification, identification, and multiple durations, in: J.J. Heckman and E. Leamer, eds. *Handbook of Econometrics, Volume V* (North-Holland, Amsterdam).

van den Berg, G.J., Lindeboom, M., Portrait, F. (2006), Economic conditions early in life and individual mortality, *American Economic Review* 96, 290–302.

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

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

van den Berg, G.J., Gupta, S., Portrait, F. (2010), Do Children Affect Life Expectancy? A Joint Study of Early life Conditions, Fertility and Mortality, *Unpublished Working Paper.*

Cawley, J., Choi, A. (2018), Health Disparities Across Education: The Role of Differential Reporting Error, *Health Economics*, 27(3): e1-e29.

Currie, J., Schwandt, H. (2016), Mortality inequality: the good news from a county-level approach. *The Journal of Economic Perspectives* 30(2), 29–52.

Cutler, D.M, Huang, W., LLeras-Muney, A. (2015), When does Education Matter for Health? The Protective Effect of Education for Cohorts Graduating in Bad Times, *Social Science and Medicine* 127, 63–73.

Cutler, D.M., Huang, W. and Lleras-Muney, A., 2016. Economic conditions and mortality: evidence from 200 years of data. *National Bureau of Economic Research*, (No. w22690).

Danmarks Statistik (1914), *Eksamensskolerne 1912* (Danmarks Statistik, Copenhagen).

Fletcher, J.M., Lehrer S.F. (2009), The Effects of Adolescent Health on Educational Outcomes: Causal Evidence Using Genetic Lotteries between Siblings, *Forum for Health Economics & Policy* 12(2), 1558–9544.

Fritze, T., Doblhammer, G. and van den Berg, G.J. (2014), Can individual conditions during childhood mediate or moderate the long-term cognitive effects of poor economic environments at birth?, *Social Science and Medicine* 119, 240–248.

Galama, T.J. and Van Kippersluis, H. (2019), A theory of socio-economic disparities in health over the life cycle, *The Economic Journal*, 129(617), 338–374.

Gilman, S.E. and Loucks, E.B (2014), Another casualty of sibling fixed-effects analysis of education and health: An informative null, or null information?, *Social Science and Medicine*, 118, 191–193.

Grossman, M. (2006), Chapter 10 Education and Nonmarket Outcomes, *Handbook of the Economics of Education* 1, 577–633.

Grossman, M. (2015), The Relationship between Health and Schooling: What's New?, *Nordic Journal of Health Economics*, 3(1), 1–7.

Halpern-Manners, A., Helgertz, J., Warren, J.R. and Roberts, E., (2020), The effects of education on mortality: Evidence from linked US Census and administrative mortality data, *Demography*, 57(4), 1513–1541.

Hernández-Quevedo, C., Jones, A. M., Rice, N. (2005), Reporting bias and heterogeneity in selfassessed health. Evidence from the British Household Panel Survey, *Health, Econometrics and Data Group (HEDG) Working Papers.*

Kaufman, J.S. and Glymour, M.M. (2011), Splitting the differences: problems in using twin controls to study the effects of BMI on mortality, *Epidemiology*, 22(1), 104–106.

Lleras-Muney, A. (2005), The Relationship Between Education and Adult Mortality in the United States, *Review of Economic Studies* 72(1), 189–221.

Loudon, I. (2000), Maternal mortality in the past and its relevance to developing countries today, *American Journal of Clinical Nutrition* 72(1), 241–246.

Ly, C.Y., Sung, F.C. (1999), A review of the healthy worker effect in occupational epidemiology, *Occupational Medicine* 49(4), 225–229.

Madsen, M., Andersen, A.N., Christensen, K., Andersen, P.K., Osler, M. (2010), Does Educational Status Impact Adult Mortality in Denmark? A Twin Approach, *American Journal of Epidemiology* 172(2), 225–234.

Malamud O, Mitrut A, Pop-Eleches C. (2021) The effect of education on mortality and health: Evidence from a schooling expansion in Romania. *Journal of Human Resources*, 1118–9863R2 .

Oreopoulos, P. (2006), Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter, *The American Economic Review* 96(1), 152–175.

Ridder, G., Tunalı, I. (1999), Stratified partial likelihood estimation, *Journal of Econometrics* 92(2), 193–232.

Schmidt, A. (2005), *Undervisningsministeriets Historie 1848–2004 (History of the Ministry of Education 1848–2004; in Danish)* (Ministry of Education, Copenhagen).

Skovgaard-Petersen, V. (1967), Den politiske drøftelse af forbindelsen mellem almueskolen og den lærde skole, *Årbog for dansk skolehistorie* 1967, 85–110.

Skytthe, A., Kyvik, K., Holm, N.V., Vaupel, J.W., Christensen, K. (2002), The Danish Twin Registry: 127 Birth Cohorts of Twins, *Twin Research* 5, 352–357.

Wienke, A. (2003), Frailty models, *MPIDR Working Papers, Max Planck Institute for Demographic Research, Rostock, Germany* WP-2003-032.

Yashin, A.I., Vaupel, J.W., Iachine, I.A. (1995), Correlated individual frailty: An advantageous approach to survival analysis of bivariate data, *Mathematical Population Studies: An International Journal of Mathematical Demography* 5(2), 145–159.

Tables: Descriptives and Estimation Results

	All twins	Same-sex twins	Monozygotic twins
Emigrated	1.56%	2.51%	3.81%
Dead	62.07%	78.82%	95.63%
Unknown	36.37%	18.66%	0.56%

Table 2: *Descriptive statistics* Number of twins (in percent), per (last known) status.

	Exposed to reform	Not exposed to reform
With exam	3.76%	2.97%
Without exam	96.24%	97.03%
Copenhagen	13.60%	12.94%
Town	20.61%	17.71%
Rural	65.80%	69.35%
Male	50.18%	49.47%
Female	49.82%	50.53%
Monozygotic (MZ)	15.50%	15.89%
DZ (same-sex)	25.29%	24.38%
Uncertain	-	0.04%
Unknown	18.07 %	19.65%
Same Sex	58.58%	59.95%
Different Sex	41.15%	40.05%

Table 3: *Descriptive statistics* Personal characteristics, by exposure to reform. Descriptive statistics on individual and twin pair characteristics and conditions at birth, split according to being eligible for the reform.

	Overall Sample	Exam	No exam
Copenhagen	13.29%	6.66%	93.34%
Town	19.26%	5.67%	94.33%
Rural	67.45%	2.15%	97.85%
Female	50.15%	2.2%	97.8%
Male	49.85%	4.58%	95.42%
Monozygotic (MZ)	15.68%	5.04%	94.96%
DZ (same-sex)	24.86%	3.81%	96.19%
Uncertain	0.02%	0.02%	0
Unknown	18.81%	19.23%	6.90%
Same Sex	59.37%	3.31%	96.69%
Different Sex	40.63%	3.50%	96.50%

Table 4: *Descriptive statistics* on twin pair characteristics and conditions at birth, tabulated by exam status.

	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916
1888	3	4	2									
1889	2	6	1					1	1			1
1890	8	8	10	4	2	0	3	1				
1891	0	0	4	5	2	7	2	1	1	0	0	0
1892	0	0	0	1	8	4	2	1	0	0	0	0
1893	0	0	0	0	7	4	3	4	0	0	1	0
1894	0	0	0	0	0	4	4	7	1	2	1	0
1895	0	0	0	0	0	1	7	13	6	3	2	0
1896	0	0	0	0	0	0	0	8	4	8	0	2
1897	0	0	0	0	0	0	0	2	10	9	4	3
1898	0	0	0	0	0	0	0	0	0	11	9	2
1899	0	0	0	0	0	0	0	0	0	0	5	0
1900	0	0	0	0	0	0	0	0	0	0	1	0

Table 5: *Descriptive statistics* Absolute number of twins in our sample that took an exam, tabulated by year of birth and the year of the exam.

Year	Number of students passing the "Studentereksamen"	Women	Men
1907	479 (+ 3.68%)	NA	NA
1908	518 (+8.14%)	78	440
1909	**516 (-0.386%)**	**54 (-30.77%)**	**462 (+4.00%)**
1910	**671 (+30.04%)**	**108 (+ 100.00%)**	**563 (+21.86%)**
1911	764 (+13.86%)	154 (+42.59%	610 (+8.35%)
1912	799 (+ 4.58%)	155 (+0.649%)	644 (+5.57%)

Table 6: *Descriptive Statistics* Aggregate number of students who passed the Studentereksamen in the years around the reform, 1910 was the first year under the new regime. The number in parentheses shows the percentage increase from the previous year.

	(1) Exam	(2) Exam	(3) Exam	(4) Exam
Reform	0.00815*	0.00712	0.00713	0.00209
	(0.00454)	(0.00540)	(0.00583)	(0.00689)
Controls	No	Yes	No	Yes
Only Same Sex	No	No	Yes	Yes
F-Statistic	3.22	11.41	1.50	8.93
Observations	6,138	6,138	3,701	3,701

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

Table 7: First stage regression of exam on reform using the full sample ((1) and (2)) and restricting to same sex only twin pairs ((3) and (4)), with ((2) and (4)) and without ((1) and (3)) additional controls.

	Full Sample	Females	Males
Exam	-0.38^{**}	-0.31	-0.50^{**}
	(0.17)	(0.23)	(0.25)
Gender	0.27^{***}	–	–
	(0.05)		
CPH	0.26^{***}	0.11	0.45^{***}
	(0.08)	(0.11)	(0.11)
Town	0.08	0.14	-0.01
	(0.07)	(0.10)	(0.10)
Logage	-0.03	-0.05	-0.02
	(0.06)	(0.08)	(0.08)
Spring	0.05	-0.06	0.16
	(0.07)	(0.10)	(0.10)
Fall	-0.04	-0.14	0.09
	(0.07)	(0.10)	(0.10)
Winter	0.05	0.01	0.08
	(0.07)	(0.10)	(0.10)
Monozygotic	-0.07	-0.07	-0.09
	(0.05)	(0.07)	(0.07)
Recession at Birth	0.02	-0.00	0.02
	(0.05)	(0.08)	(0.08)
Recession at Graduation	0.01	-0.01	0.04
	(0.05)	(0.08)	(0.08)
Recession*exam	0.43	0.46	0.42
	(0.24)	(0.38)	(0.31)
B	2.415	2.413	2.426
	(0.043442)	(0.062748)	(0.060808)
ρ_{dz}	0.827	0.918	0.708
ρ_{mz}	0.955	1.00	0.870
γ	16.41	15.82	17.50
Constant	-6.26^{***}	-6.15^{***}	-6.08^{***}
	(0.17)	(0.24)	(0.25)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 8: Regression Results from the full model with *exam* as the education covariate.

	Full Sample	Females	Males
Reform	-0.10	0.11	-0.28^{*}
	(0.11)	(0.16)	(0.15)
Gender	0.27^{***}	–	–
	(0.05)		
CPH	0.23^{***}	0.09	0.38^{***}
	(0.07)	(0.10)	(0.11)
Town	0.08	0.13	-0.01
	(0.07)	(0.10)	(0.10)
Logage	0.05	-0.15	0.21
	(0.10)	(0.15)	(0.15)
Spring	0.04	-0.07	0.15
	(0.07)	(0.10)	(0.10)
Fall	-0.04	-0.15	0.09
	(0.07)	(0.10)	(0.10)
Winter	0.05	0.01	0.08
	(0.07)	(0.10)	(0.10)
Monozygotic	-0.08	-0.07	-0.11
	(0.05)	(0.07)	(0.07)
Recession at Birth	0.03	-0.01	0.06
	(0.05)	(0.08)	(0.08)
Recession at Graduation	0.00	0.04	-0.01
	(0.06)	(0.09)	(0.09)
B	2.4111	2.4136	2.4171
	(0.043394)	(0.062885)	(0.060654)
ρ_{dz}	0.827	0.909	0.715
ρ_{mz}	0.962	1.000	0.894
γ	16.59	17.07	17.65
Constant	-6.32^{***}	-6.05^{***}	-6.30^{***}
	(0.19)	(0.26)	(0.27)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 9: Regression Results from the correlated frailty model, estimating the effect of *reform* on mortality.

Exam	-1.95	(1.07)
Exam $\times$ recession at birth	2.55	(1.18)

Table 10: *Stratified partial likelihood estimation results* (s.e. in parentheses)

A Technical Appendix

A.1 Derivation of the correlated frailty model

The likelihood function for the correlated frailty model for all observation pairs is given by:

$$
\begin{aligned}
L(t_1, t_2) &= \delta_1\delta_2\bar{F}_{t_1,t_2}(t_1, t_2) + \delta_1(1-\delta_2)(\bar{F}_{t_2}(t_1, t_2) + \bar{F}_{t_1}) \\
&+ \delta_2(1-\delta_1)(\bar{F}_{t_1}(t_1, t_2) + \bar{F}_{t_2}) + (1-\delta_1)(1-\delta_2)(\bar{F}(t_1, t_2))
\end{aligned} \tag{8}
$$

where $\delta_i = 0$ if observation i is uncensored and 1 otherwise. For two uncensored observations ($\delta_1 = \delta_2 = 1$) the bivariate probability density function is given by the following expression:

$$
\begin{aligned}
f(t_1, t_2) &= \frac{(1-\rho)^2\bar{F}(t_1)^{-\rho}\bar{F}(t_2)^{-\rho}f(t_1)f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_1)^{-\sigma^2} - 1)^{\rho/\sigma^2}} + \frac{(1-\rho)\rho\bar{F}(t_1)^{-\rho}\bar{F}(t_2)^{-\rho-\sigma^2}f(t_1)f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+1}} \\
&+ \frac{(1-\rho)\rho\bar{F}(t_1)^{-\rho-\sigma^2}\bar{F}(t_2)^{-\rho}f(t_1)f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+1}} + \frac{\rho(\rho+\sigma^2)\bar{F}(t_1)^{-\rho-\sigma^2}\bar{F}(t_2)^{-\rho-\sigma^2}f(t_1)f(t_2)}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{(\rho/\sigma^2)+2}}
\end{aligned} \tag{9}
$$

For $\delta_1 = \delta_2 = 0$ the bivariate survivor function has the following form:

$$
\bar{F}(t_1, t_2) = \frac{\bar{F}(t_1)^{1-\rho}\bar{F}(t_2)^{1-\rho}}{(\bar{F}(t_1)^{-\sigma^2} + \bar{F}(t_2)^{-\sigma^2} - 1)^{\rho/\sigma^2}} \tag{10}
$$

For $\delta_1 = 1, \delta_2 = 0$ the pairwise contribution to the likelihood function is

$$
\begin{aligned}
\partial\bar{F}(t_1, t_2)/\partial t_1 &= (1-\rho)\bar{F}_1^{-\rho}\bar{F}_2^{1-\rho}(-f_1)(\bar{F}_1^{-\sigma^2} + \bar{F}_2^{-\sigma^2} - 1)^{\frac{-\rho}{\sigma^2}} \\
&- \bar{F}_1^{1-\rho}\bar{F}_2^{1-\rho}(\bar{F}_1^{-\sigma^2} + \bar{F}_2^{-\sigma^2} - 1)^{(-1-\frac{\rho}{\sigma^2})}\bar{F}_1^{-\sigma^2-1}f_1
\end{aligned} \tag{11}
$$

For $\delta_1 = 0, \delta_2 = 1$ the pairwise contribution to the likelihood function is

$$
\begin{aligned}
\partial\bar{F}(t_1, t_2)/\partial t_2 &= (1-\rho)\bar{F}_2^{-\rho}\bar{F}_1^{1-\rho}(-f_2)(\bar{F}_2^{-\sigma^2} + \bar{F}_1^{-\sigma^2} - 1)^{\frac{-\rho}{\sigma^2}} \\
&- \bar{F}_2^{1-\rho}\bar{F}_1^{1-\rho}(\bar{F}_2^{-\sigma^2} + \bar{F}_1^{-\sigma^2} - 1)^{(-1-\frac{\rho}{\sigma^2})}\bar{F}_2^{-\sigma^2-1}f_2
\end{aligned} \tag{12}
$$

If we assume the Gompertz function for the baseline hazard then, with $\gamma := \frac{1}{\sigma^2}$,

- $\bar{F}(t_i) = (1 + e^{x_i}\gamma^{-1}B^{-1}(e^{Bt_1} - 1))^{-\gamma}$
- $f(t_i) = e^{x_i'\beta}e^{Bt_i}(1 + \gamma^{-1}B^{-1}(e^{Bt_1} - 1))^{-\gamma-1}$

A.2 Illustration of dynamic selection

To illustrate the importance of controlling for unobserved heterogeneity, consider the very simple case of a constant mortality hazard, a binary treatment T, and two different groups with associated frailty u_h and u_l that are unobserved. Individuals with unobserved type u_h have ex-ante a higher probability of dying. The treatment corresponds to a constant shift in the mortality hazard in each period, so that in this simplified example we have the following mortality probabilities in each period, represented in Table (11). In this case, the treatment (for example education) cuts the probability of dying in half for all frailty groups, resulting in a constant shift in the relative risk of dying within each frailty group. The frailty group, in this case, may represent something unobservable or possibly observable but not included in the analysis. Notice that the frailty status is independent of the treatment status, in fact, in the simulations below we evenly distribute the observations, so that there is the same number of observations (initially) in each group with $N = 4000$. This set-up corresponds to analyzing life-tables where we observe the number of people at risk in both treatment and control group and subsequent deaths. Since the frailty type $u.$ is unobserved, this means that we sum over all individuals of both types and end up with $N_T = 2000$ observations in the treatment group and $N_C = 2000$ observations in the control group. To evaluate the effect of the treatment, we will focus on the reduction in mortality risk as a result of being assigned to either the treatment or control group by dividing the number of deaths in each group per period by the number of individuals still alive. Given the time constant probability of dying and the constant effect of the treatment (cutting the probability of dying in half in each period), we should expect a constant shift in the hazard of dying.

We plot the results in Figure 4. The top left panel depicts the risk of dying over all 60 periods for the control group (black dashed line) and the treatment group (red solid line). While it initially appears as though the treatment has a large effect on mortality, the effect, i.e. the difference between the two curves seems to shrink over time (see also the difference is depicted in the top right panel). How is that possible when we *know* that the treatment has a constant effect on survival probabilities? Pulling back the curtain, we can also calculate the risk for treatment and control group within each unobserved frailty group and notice that within groups, the treatment does have a constant effect on the risk of dying (depicted in the lower-left panel). When we examine the group composition over time, depicted on the right panel, we see stark differences between treatment and control group: the figure shows the ratio of high frailty types to low frailty types and is initially equal to one, per our assignment and subsequently falls much quicker to zero for the control group. This is because the treatment affects both high and low frailty types and thus high frailty observations are kept around longer in the treatment group, affecting our estimates. In conclusion: the above example demonstrates that we would be making two mistakes when calculating dynamic treatment effects (for an exogenous treatment) in this manner: (1) We would conclude that the risk of dying decreases over time when it is, in fact, constant and (2) that the size of the treatment effect decreases over time from initially cutting mortality risk in half to only decreasing the mortality risk by a little more than one third when the treatment consistently reduces mortality risk over all periods.

Knowing the above, one could ask: what happens if the researcher ignores the dynamic aspect and just compares average age-at-death between treatment and control group? Would this provide the correct answer? Unfortunately, the estimates of the treatment effect will also be downward biased in the presence of left or right censoring: In the case of our example, this means that we would ex-ante assume that, because the treatment reduces mortality hazard by 1/2 in each period, observations in the treatment group should live twice as long as observations in the control group, i.e. the ratio of the average age-at-death between control group and treatment group should be around 0.5. However: this is only true if there is no censoring: right censoring refers to the phenomenon that we sometimes only observe cohorts until say age 80. Left truncation means that we do not observe units the moment they come under risk, but only later, when selective mortality has already occurred. A realistic scenario of this would be that we evaluate the life-tables of cohorts for an educational reform that occurred, for example in 1920, and reliable life-tables are only available for the relevant cohorts starting in 1950.

Figure 5 illustrates this by showing the ratio of age-at-death estimates at different censoring and truncation times.[12] We increase the number of periods to 400 to ensure that almost everyone is dead and then (secularly) increase the amount of right censoring (black dashed line) and left truncation (solid blue line). We can see that if there is no censoring then the ratio of age-at-death for the control group and treatment group is almost 1/2 (which corresponds to the true effect). However, as soon as we introduce either right censoring or left truncation, this ratio goes to one, meaning that we would conclude that the treatment has a much smaller effect than it does.

[12] A censoring time of 20 corresponds in the case of left truncation of observations entering the sample in period 20, for right censoring it means observations are only observed until period 380.

	Control Group		Treatment Group	
Frailty Group	u_l	u_h	u_l	u_h
Mortality Hazard	0.02	0.08	0.01	0.04

Table 11: Mortality hazards for treatment groups and control groups

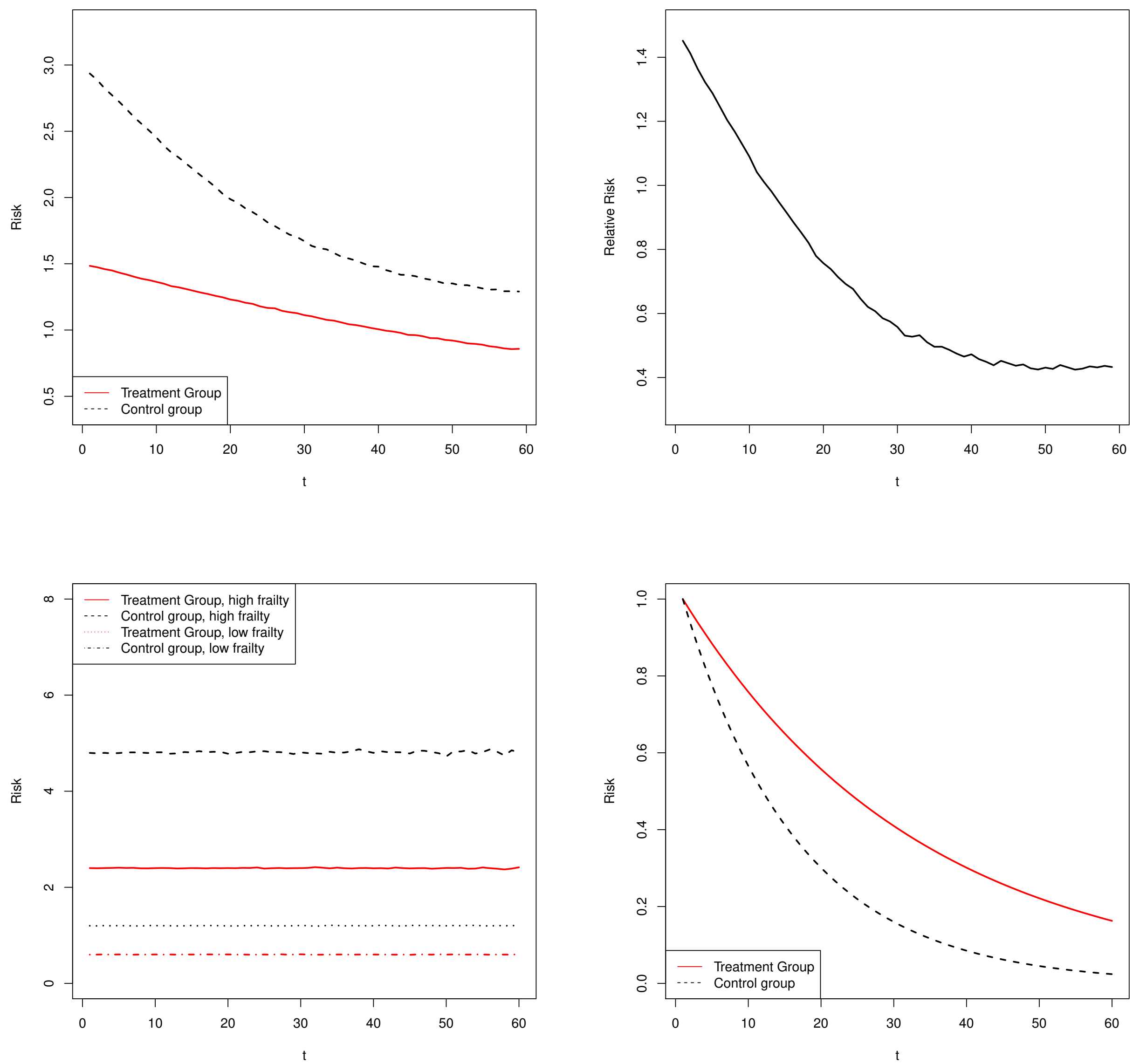

Figure 4: Top left panel depicts the risk of dying for the treatment group (red solid line) and the control group (black dashed line).

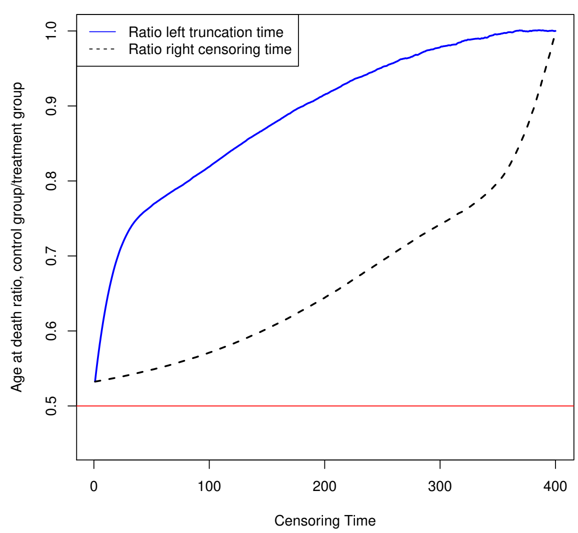

Figure 5: Average age at death in the treatment group divided by average age at death of the control group for different degrees of right censoring and left truncation.

B Data Collection

B.1 Education and Mortality Data

To document our data collection process for the educational data, we have included an example page of our scans and another page showing the result of the text recognition process. The first column shows the name of the student, the second column the parish of birth and the date of birth, the third column shows the father's occupation and the fifth column shows the admission date. As can be seen from the processed page (7), the names and the parish of birth were recognized with a very high degree of accuracy. The exact birth dates were added by hand later and some random checks were made as well. We then matched this retrieved data using a data base program and pre-filtered by the first three letters of the last name (including spelling variations) and the year of birth. We then manually matched on birth date and exact name, cases that were questionable were later double checked in the original scans.

Figure 6:

40 De højere Almenskoler 1909—10.

Eksaminandernes Navne.	Fødested, Fødselsaar og -dag.	Faderens Stilling.	Naar optaget i Skolen og i hvilken Klasse?
Cramer, Hans Billeskov Jansen	Trælund $^{27}/_{5}$ 91	Gaardejer.	Aug. 07 1 G.
Gøtzsche, Carl Viggo	København $^{26}/_{12}$ 92	Sognepræst.	— 07 1 —
Hvilsom, Knud Emil Frederik	Ans $^{19}/_{8}$ 91	Læge.	— 03 1 M.
Jensen, Jeppe	Mønsted $^{21}/_{2}$ 90	Skovløber.	Sept. 04 2 —
Larsen, Laurits Kristian	Hundborg $^{11}/_{3}$ 88	Husmand.	Okt. 07 1 G.
Nielsen, Jensen	Ejerslev $^{20}/_{1}$ 91	Gaardejer.	Aug. 07 1 —
Olsen, Elise Dorothea	Viborg $^{5}/_{4}$ 91	Rektor.	— 03 1 M.
Pedersen, Jens Carl	Grødde $^{3}/_{6}$ 89	Husmand.	— 07 1 G.
Petersen, Knud Honoré	Viborg $^{18}/_{11}$ 92	Oversergent.	Apr. 06 3 M.
Rossen, Ejnar	Bjerringbro $^{17}/_{8}$ 91	Læge.	Aug. 05 3 —
Siersted, Hans Christian	Herning $^{15}/_{2}$ 93	Herredsfoged.	— 07 1 G.
Den matematisk-naturvidenskabelige Linie.			
Jørgensen, Jørgen Hansen Viggo	Terpling $^{8}/_{2}$ 92	Realskolebest.	Aug. 07 1 G.
Obbekjær, Andreas Thorup	Helligkilde $^{29}/_{12}$ 90	Lærer.	— 07 1 —
Rendtorff, Theodor Herman Otto	Hemmestrup $^{15}/_{6}$ 92	Proprietær.	— 03 1 M.
Sørensen, Christen Nordentoft	Thaliasminde $^{29}/_{2}$ 92	Gaardejer.	— 07 1 G.
Aarhus Katedralskole.			
Den klassisk-sproglige Linie.			
Børup, Marinus	Aarhus $^{27}/_{6}$ 91	Skrædder.	Aug. 04 2 M.
Sejr, Emanuel Jensen	Nølev $^{29}/_{7}$ 91	Lærer.	— 04 2 —
Den nysproglige Linie.			
Andresen, Gerda Johanne	Aarhus $^{4}/_{2}$ 91	Sagfører.	Aug. 04 2 M.
Jensen, Thomas Agner	Ry $^{5}/_{9}$ 91	Lærer.	— 07 1 G.
Paasgaard, Marie Elisabeth	Aarhus $^{14}/_{8}$ 90	Overpolitibetj.	— 03 1 M.
Pøckel, Emmy	— $^{24}/_{8}$ 90	Fuldmægtig.	— 04 2 M.
Rud, Ejnar Jensen	Viby $^{7}/_{2}$ 92	Redaktør.	— 04 2 —
Sejersen, Ellen	København $^{23}/_{10}$ 92	Købmand.	— 04 2 —
Smith, Karen Wiggers	Holbæk $^{16}/_{2}$ 93	Konsul.	— 07 1 G.
Sørensen, Ingeborg	Jelling $^{22}/_{10}$ 91	Biskop.	— 07 1 —
Thomsen, Holger de Fine	Hjørring $^{6}/_{6}$ 92	Herredsfoged.	— 03 1 M.
Den matematisk naturvidenskabelige Linie.			
Efsen, Axel Valdemar	Silkeborg $^{1}/_{1}$ 93	Fabrikant.	Aug. 07 1 G.
Fricke, Hugo	Aarhus $^{15}/_{8}$ 92	Stabssergent.	— 07 1 —
Hammelev, Kaj	København $^{2}/_{1}$ 92	Oberst.	Okt. 07 1 —
Kühnel, Poul Oscar	Aarhus $^{24}/_{2}$ 92	Arkitekt.	Juli 03 1 M.
Lauesen, Henrik Møller Langkilde	Skanderborg $^{19}/_{5}$ 92	Overretssagf.	Aug. 06 4 —
Nielsen, Marinus Michael	Aarhus $^{16}/_{10}$ 91	Købmand.	— 04 2 —
Qvist, Peter Martin	København $^{17}/_{3}$ 92	Laboratorieforst.	— 03 1 —
Sand, Ernst Frederik	Aarhus $^{3}/_{3}$ 92	Fabrikbestyrer.	— 03 1 —
Ribe Katedralskole.			
Den klassisk-sproglige Linie.			
Binzer, August Reinhardt	Godthaab, Grønland $^{28}/_{12}$ 90	Distriktslæge.	Aug. 02 1 st.
Dela, Hans Peter	Ribe $^{14}/_{1}$ 92	Avlsbruger.	— 03 1 —
Fogh, Knud Adolf Thorvald Amelius	Holstebro $^{16}/_{3}$ 92	Amtsforvalter.	— 03 1 M.

Example page, high quality scan

Figure 7:

Eksaminandernes Navne.	Fødested, Fødselsaar og -dag.	Faderens Stilling.	Naar optaget i Skolen og i hvilken Klasse?
Cramer, Hans Billeskov Jansen	Trælund 27/3 91	Gaardejer.	Aug. 07 1 G.
Gøtzsche, Carl Viggo	København 21/i, 92	Sognepræst.	— 07 1 —
Hvilsom, Knud Emil Frederik	Ans 19/s 91	Læge.	— 03 1 M.
Jensen, Jeppe	Mønsted 21/2 90	Skovløber.	Sept. 04 2 —
Larsen, Laurits Kristian	Hundborg "/e 88	Husmand.	Okt. 07 1 G.
Nielsen, Jensen	Ejerslev 20/x 91	Gaardejer.	Aug. 07 1 —
Olsen, Elise Dorothea	Viborg 7, 91	Rektor.	— 03 1 M.
Pedersen, Jens Carl	Grødde 3/e 89	Husmand.	— 07 1 G.
Petersen, Knud Honoré	Viborg 18/n 92	Oversergent.	Apr. 06 3 M.
Eossen, Ejnar	Bjerringbro 17/8 91	Læge.	Aug. 05 3 —
Siersted, Hans Christian	Herning »/, 93	Herredsfoged	— 07 1 G.
Den matematisk-natur-videnskabelige Linie.			
Jørgensen, Jørgen Hansen Viggo	Terpling 8/2 92	Realskolebest.	Aug. 07 1 G.
Obbekjær, Andreas Thorup	Helligkilde 29/12 90	Lærer.	07 1 -
Rendtorff, Theodor Herman Otto	Hemmestrup 13/6 92	Proprietær.	03 1 M.
Sørensen, Christen K ordentoft.,. -	Thaliasminde 29/2 92	Gaardejer.	07 1 G
Aarhus Katedralskole.			
Den klassisksproglige Linie.			
Børup, Marinus	Aarhus 27/6 91	Skrædder.	Aug. 04 2 M.
Sejr, Emanuel Jensen	Nølev 29/7 91	Lærer.	— 04 2 -
Den nysproglige Linie.			
Andresen, Gerda Johanne	Aarhus 7,	Sagfører.	Aug. 04 2 M.
Jensen, Thomas Agner	Ry	Lærer.	— 07 1 G.
Paasgaard, Marie Elisabeth.,	Aarhus 14/8 90	Overpolitibetj.	— 03 1 M.
Pøckel, Emmy	17, 90	Fuldmægtig.	— 04 2 M.
Bud, Ejnar Jensen	Viby 7, 92	Redaktør.	_ 04 2 —
Sejersen, Ellen	København 8/io 92	Købmand.	— 04 2 —
Smith, Karen Wiggers	Holbæk 16/2 93	Konsul.	— 07 1 G.
Sørensen, Ingeborg	Jelling 7io 91	Biskop.	— 07 1 —
Thomsen, Holger de Fine	Hjørring % 92	Herredsfoged.	— 03 1 M.
Den matematisk naturvidenskabelige Linie.			
Efsen, Axel Valdemar	Silkeborg 7i 93	Fabrikant.	Aug. 07 1 G.
Fricke, Hugo	Aarhus 13/8 92	Stabssergent.	— 07 1 —
Hammelev, Kaj	København 2/t 92	Oberst.	Okt. 07 1 —
Kuhnel, Poul Oscar	Aarhus 24/2 92	Arkitekt.	Juli 03 1 M.
Lauesen, Henrik Møller Langkilde	Skanderborg 19/3 92	Overretssagf.	Aug. 06 4 —
Melsen, Marinus Michael	Aarhus 16/10 91	Købmand.	— 04 2 —
Qvist, Peter Martin	København 17/3 92	Laboratorieforst.	— 03 1 —
Sand, Ernst Frederik		Fabrikbestyrer.	— 03 1 —
Ribe Katedralskole.			
Den klassisksproglige Linie.			
Binzer, August Keinhardt	Godthaab, Grønland 28/12 90	Distriktslæge.	Aug. 02 1 st.
Dela, Hans Peter.	Ribe 14/i 92	Avlsbruger.	— 03 1 —
Fogh, Knud Adolf Thorvald Amellus	Holstebro 16/3 92	Amtsforvalter.	— 03 1 M.

Example page, processed with text recognition

B.2 Occupations and Socio-Economic Group Classification

The classification scheme that we use for understanding the make-up of the socio-economic groups is derived from the swedish classification schedule for socio-economic groups ("Klassificeringsschema för socio-ekonomiska grupper") from Swedens "Folk-och Bostadsräkningen". The classification is based on the occupation ("Yrkeskod") and the position (entrepreneur/self-employed, employee, worker). There are 10 different groups, though we do not report the results on group 10 ("no identified occupation"). We use father's occupation at the time of the exam as shown in the third column in Figure 6 and classify it according to the occupation codes. For example, "Gaardejer" would have occupation code "401" and the associated social class code is "1".

1. Entrepreneurs/ landowners in agriculture (*Företagare inom jordbruk, skogsbrukm. m.*).
2. Employees/workers in agriculture (*Anställda inom jordbruk, skogsbrukm. m.*).
3. Entrepreneurs within industry, trade, logistics, service occupations (*Företagare inom industri–, handels-,transport- och serviceyrken*).
4. Self–employed professionals (Physicians, Lawyers,...) (*Företagare inom fria yrken (läkare,advokater m. f l .).*).
5. Employed executives *Företagsledare (anställda).*
6. Employed professionals *Anställda inom tekniska, humanistiska,kontorstekniska och kommersiellayrken m. fl..*
7. Workers in industry and logistics (*Anställda inom huvudsakligen industri-och transportyrken*).
8. Service Workers (*Anställda inom vissa serviceyrken(hembiträden, serveringspersonal m.fl.)*).
9. Military personnel (all levels) (*Militärer*).
10. Personer med ej identifierbarayrken.

Figure 8:

1960 års folkräkning
21.2.61.

YRKESKOD

0.. Naturvetenskapligt, tekniskt samt socialvetenskapligt, *humanistiskt och konstnärligt* arbete

00. Tekniskt arbete

001 Arkitekter, ingenjörer och tekniker med *byggnads- och anläggningstekniskt* arbete
002 Ingenjörer och tekniker med elkraft-tekn. och teletekn. arbete
003 Ingenjörer och tekniker med mekaniskt arbete
004 Ingenjörer och tekniker med kemitekn. arbete
005 Ingenjörer och tekniker med gruvtekn. och metallurgiskt arbete
006 Ingenjörer och tekniker inom andra *tekn.* verksamhetsområden
007 Lantmätare, mätningstekn., kartografer
008 Tekniska biträden
009 Ej specificerbar uppgift

01. Kemiskt och fysikaliskt arbete

011 Kemister
012 Fysiker
013 Geologer, meteorologer m. fl.
019 *Ej spe*cificerbar *uppgift*

02. Biologiskt arbete

021 Veterinärer
022 Biologer
*023 Jordbruks- och trädgårdsforskare/-rådgi-*vare
024 Skogsbruksforskare, skogsbruksrådgiv.
029 Ej specificerbar uppgift

03. Medicinskt arbete

031 Läkare
032 Tandläkare
039 Ej specificerbar uppgift

04. Hälso- och sjukvårdsarbete

040 Sjuksköterskor
041 Barnmorskor
042 Skötare inom mentalvård
043 Sjukvardsbiträden
044 Tandsköterskor
045 Röntgenbehandlingsbiträden, laboratriser m. fl.
046 Farmaceuter
047 Sjukgymnaster, massörer m. fl.
048 Hälsovårdsinspektörer m. fl.
049 Ej specificerbar uppgift

05. Pedagogiskt arbete

050 Skolledare
051 Universitets- och högskolelärare
052 Lärare i läroämnen
053 Klasslärare
054 Lärare i övningsämnen
055 Yrkeslärare
056 Förskollärare
057 Utbildningskonsulenter m. fl.
058 Övrigt pedagogiskt arbete
059 Ej specificerbar uppgift

06. Religiöst arbete

061 Präster och predikanter
068 Övrigt religiöst arbete
069 Ej specificerbar uppgift

07. Juridiskt arbete

071 Domstolsjurister
072 Åklagare och högre polistjänstemän
073 Praktiserande jurister m. fl.
074 Juridiska ombudsmän
078 Övrigt juridiskt arbete
079 Ej specificerbar uppgift

08. Litterärt och konstnärligt arbete

081 Bildkonstnärer
082 *Formgivare*
083 Dekoratörer
084 Författare
085 Journalister, förlagsredaktörer
086 Scenkonstnärer
087 Musiker
088 Övrigt litterärt och konstnärligt arbete
089 Ej specificerbar uppgift

09. Övrigt tekniskt, naturvetenskapligt m. m. arbete

091 Revisions- och redovisningsexperter
092 Socialtjänstemän
093 Bibliotekarier, arkivarier, museitjänstemän
094 Ekonomer, statistiker
095 Psykologer, personalmän
098 Övrigt hithörande arbete
099 Ej specificerbar uppgift

1.. *Administrativt arbete*

10. Allmänt samhällsadministrativt arbete

101 Allmänt samhällsadministrativt arbete

11. Företagsadministrativt samt annat tekniskt och ekonomiskt administrativt arbete

111 Företagsledare
118 Övriga *företagsadministratörer* och administratörer för speciella funktioner (såväl enskild som statlig och kommunal tjänst)
119 Ej specificerbar uppgift

2.. Kameralt och kontorstekniskt arbete

20. Bokförings- och kassaarbete

201 Bokförare och kontorskassörer
203 Bankkassörer
204 Butiks- och restaurangkassörer
208 Inkasserare m. fl.
209 Ej specificerbar uppgift

29. Stenografi- och maskinskrivnings- samt annat kontorsarbete

290 Sekreterare, stenografer, maskinskrivare
291 *Datamaskinoperatörer*
292 Banktjänstemän
293 Resebyråtjänstemän
294 Speditörer, skeppsklarerare m. fl.
295 Förvaltare av fast och lös egendom
296 Försäkringstjänstemän
297 Sjukkassetjänstemän
298 Specialkontorister m. fl.
299 Ej specificerbar uppgift

3.. *Kommersiellt arbete*

30. Parti- och detaljhandelsföretagare

301 Grosshandlare
302 Detaljhandlare
309 Ej specificerbar uppgift

31. Försäljning av fast och lös egendom, tjänster och värdebevis m. m.

311 Försäkringsförsäljare
312 Egendoms- och värdepappersmäklare
313 Reklammän
318 Värderingsmän, auktionister
319 Ej specificerbar uppgift

***32. Partiförsäljning genom kundbesök* samt agenturarbete**

321 Handelsresande, agenter

33. Övrigt kommersiellt arbete

331 Inköpare, kontorsförsäljare
332 Affärsföreståndare
333 Övrig affärspersonal
334 Ambulerande försäljare
338 Bensinförsäljare, demonstratörer m. fl.
339 Ej specificerbar uppgift

4.. Lantbruks-, skogs- och fiskeriarbete

40. Lantbruks-, skogs- och trädgårdsledning

401 Lantbrukare, skogsbrukare och trädgårdsbrukare
402 Lantbruksbefäl
403 Skogsbefäl
404 Trädgårdsbefäl
405 Husdjursuppfödare
406 Pälsdjursuppfödare
407 Renägare
409 Ej specificerbar uppgift

***41. Jordbruks- och trädgårdsarbete,* djurskötsel**

411 Lantarbetare
412 Trädgårdsarbetare
413 Husdjursskötare
414 Pälsdjursskötare
415 Renskötare
418 Övrigt jordbruks- och trädgårdsarbete m. m.
419 Ej specificerbar uppgift

42. Viltvård och jakt

421 Viltvårdare och jägare

43. Fiskeriarbete

431 Fiskare
432 Fiskodlare
439 Ej specificerbar uppgift

44. Skogsarbete

441 Skogs- och flottningsarbetare

5.. Gruv- och stenbrytningsarbete

50. Gruv- och stenbrytningsarbete
501 Gruvbrytare, bergsprängare m. fl.
502 Brunnsborrare, diamantborrare
503 Anrikningsarbetare
504 Övriga gruv- och stenbrytningsarbetare
509 Ej specificerbar uppgift

6.. Transport- och kommunikationsarbete

60. Sjöbefälsarbete
601 Nautiskt befäl
602 Lotsar
603 Maskinbefäl
609 Ej specificerbar uppgift

61. Däcks- och maskinmanskap

611 Däcks- och maskinmanskap

62. Flygarbete

621 Flygförare, flygmaskinister m. fl.

63. Lokförare, järnvägskonduktörer, trafikbiträden samt vägtrafikarbete

631 Lokförare, lokbiträden
632 Järnvägskonduktörer, trafikbiträden
633 Motorfordonsförare, spårvagnsförare
634 Kuskar
635 Cykelbud m. fl.
636 Buss- och spårvagnskonduktörer, spärrvakter
639 Ej specificerad uppgift

Figure 9:

64. Trafikledning och trafikarbetsledning

641 Hamntrafikledare
642 Flygtrafikledare, flygklarerare m. fl.
643 Trafikbefäl vid järnväg
644 Vägtrafikledare
649 Ej specificerbar uppgift

65. Post- och telekommunikation

651 Postassistenter, postexpeditörer m. fl.
652 Teleassistenter m. fl.
653 Telefonister (televerket).
654 Kontorstelefonister
655 Telegrafexpeditörer, radioexpeditörer
659 Ej specificerbar uppgift

66. Postalt och annat budarbete

661 Postiljoner
662 Expeditionsvakter, kontorsbud m. fl.
669 Ej specificerbar uppgift

67. Övrigt transport- och kommunikationsarbete

671 Fyrvaktare, sluss-, färj- och hamnvakter m. fl.
678 Banbiträden i banvaktstjänst

69. Ej specificerat transport- och kommunikationsarbete

699 Ej specificerat transport- och kommunikationsarbete

7../8.. Tillverkningsarbete

70. Textilarbete

701 Textilarbetare

71. Sömnadsarbete

711 Skräddare, ateljésömmerskor m. fl.
712 Körsnärer
713 Modister och hattmakare
714 Tapetserare
715 Tillskärare m. fl.
716 Konfektionssömmerskor
718 Övrigt sömnadsarbete
719 Ej specificerbar uppgift

72. Sko- och läderarbete

721 Skomakare
722 Skoarbetare
726 Sadelmakare, lädersömmare m. fl.
729 Ej specificerbar uppgift

73. Järnbruks-, metallverks-, smides- och gjuteriarbetare

731 Hytt- och metallugnsarbetare
732 Härdare, glödgare m. fl.
733 Varmvalsare, kallvalsare
735 Smeder
736 Gjuteriarbetare
737 Tråddragare, rördragare
738 Övrigt järnbruks-, metallverks-, smides- och gjuteriarbete
739 Ej specificerbar uppgift

74. Finmekaniskt arbete

741 Instrumentmakare (finmekaniker)
742 Urmakare
743 Optiker
744 Tandtekniker
745 Guld- och silversmeder
749 Ej specificerbar uppgift

75. Verkstads- och byggnadsmetallarb.

750 Verkstadsmekaniker (bänk- och maskinarbetare)
751 Montörer-maskinuppsättare
752 Maskin- och motorreparatörer
753 Tunnplåtslagare
754 Rörledningsarbetare
755 Svetsare, gasskärare
756 Grovplåtslagare och stålkonstruktionsarbetare
757 Metalliserare m. fl.
758 Övrigt verkstads- och metallind.arb.
759 Ej specificerbar uppgift

76. Elektroarbete

761 Installations-, drifts- och maskinelektr.
764 Tele-, radio- och TV-reparatörer
766 Telefonreparatörer-installatörer
767 Linjearbetare
768 Övrigt elektroarbete
769 Ej specificerbar uppgift

77. Träarbete

771 Byggnadsträarbetare
772 Bänk- och maskinsnickare, möbelsnick. m. fl.
774 Ram- och cirkelsågare, hyvlare m. fl.
778 Övrigt träarbete
779 Ej specificerbar uppgift

78. Målnings- och lackeringsarbete

781 Målare
782 Lackerare
789 Ej specificerbar uppgift

79. Mureri- och betongarbete m. m.

791 Murare, rappare, putsare
792 Stenmontörer
793 Byggnadsspecialarbetare, byggnadsgrovarbetare
794 Isolerare
795 Glasmästeriarbetare
798 Övrigt mureri- och betongarbete
799 Ej specificerbar uppgift

80. Grafiskt arbete

801 Typografer, litografer m. fl.
806 Bokbinderiarbetare
808 Övrigt grafiskt arbete
809 Ej specificerbar uppgift

81. Glas-, porslins-, keramik- och tegelarbete

811 Glashyttarbetare
812 Formare (keramiska produkter)
813 Ugnsskötare (glas och keramisk tillv.)
814 Dekoratörer, glaserare (glas, porslin, keramik)
818 Övrigt glas-, porslins-, keramik- och tegelarbete
819 Ej specificerbar uppgift

82. Livsmedelsarbete

821 Kvarnarbetare
822 Bagare och konditorer
823 Choklad- och sötvaruarbetare
824 Bryggeri-, vattenfabriks- och bränneriarbetare m. fl.
825 Konservarbetare
826 Slakteri- och charkuteriarbetare
827 Mejeriarbetare
828 Övrigt livsmedelsarbete
829 Ej specificerbar uppgift

83. Kemiskt- och cellulosatekniskt arb.

831 Kemiska processarbetare
834 Trämasseslipare, cellulosaarbetare
836 Pappers-, papp- och fiberplattarbetare
838 Övrigt kemiskt- och cellulosatekniskt arbete
839 Ej specificerbar uppgift

84. Tobaksarbete

841 Tobaksarbetare

85. Övrigt tillverkningsarbete

850 Korgmakeriarbetare
851 Gummivaruarbetare
852 Plastvaruarbetare
853 Garvare och skinnberedare
854 Fotolaboratoriearbetare
855 Musikinstrumentmakare
856 Stenhuggeriarbetare
857 Pappers- och emballagearbetare
858 Övrigt tillverkningsarbete

86. Grov- och diversearbete

661 Grov- och diversearbetare

87. Maskin- och motorskötsel

871 Landmaskinister
872 Kran- och traversförare
873 Riggare
874 Anläggningsmaskinförare
875 Truckförare, transportörskötare m. fl.
876 Smörjare
879 Ej specificerbar uppgift

88. Paketerings- och emballeringsarb. samt stuveri-, lager- och förrådsarb.

881 Paketerare och emballerare
882 Stuveriarbetare samt andra lastnings- och lossningsarbetare
883 Lager- och förrådsarbetare
888 Flyttkarlar m. fl.
889 Ej specificerbar uppgift

89. Ej specificerbart tillverkningsarbete

899 Ej specificerbart tillverkningsarbete

90.—94. Servicearbete

90. Civilt bevaknings- och skyddsarbete

901 Brandmän
902 Poliser
903 Tullbevakningspersonal
904 Fångvårdare
908 Övrigt bevaknings- och skyddsarbete
909 Ej specificerbar uppgift

91. Husligt arb., portierarbete m. m.

911 Ekonomiföreståndare
912 Kockar och kallskänkor
913 Köksbiträden
914 Hembiträden och barnsköterskor
915 Hemvårdarinnor m. fl.
916 Hotellportierer
917 Pursers, trafikvärdinnor m. fl.
918 Övrigt husligt arbete
919 Ej specificerbar uppgift

92. Serveringsarbete

921 Hovmästare, servitörer

93. Fastighetsskötsel, städning

931 Fastighetsarbetare m. fl.
932 Städare
933 Skorstensfejare
939 Ej specificerbar uppgift

94. Övrigt servicearbete

941 Frisörer, skönhetsvårdare m. fl.
942 Badpersonal
943 Tvättare
944 Pressare
945 Sportledare, travtränare, jockeys m. fl.
946 Fotografer
947 Begravningsbyråföreståndare m. fl.
948 Övrigt servicearbete
949 Ej specificerbar uppgift

98. Militärt arbete

981 Militärt arbete

99. Personer med ej identifierbara yrken eller med ej angiven yrkestillhörighet

999 Personer med ej identifierbara yrken el. med ej angiven yrkestillhörighet

Figure 10:

Klassificeringsschema för socio-ekonomiska grupper *Bilaga 10*

Schedule of socio-economic groups

Medhjälpande i företagarhushållen har inplacerats i samma grupp som företagaren. Ej förvärvsarbetande familjemedlemmar (hustrur samt barn under 16 år) har erhållit huvudmannens socio-ekonomiska grupp.

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
001	4	6	-
002	4	6	-
003	4	6	-
004	4	6	-
005	4	6	-
006	4	6	-
007	4	6	-
008	-	6	-
009	4	6	-
011	4	6	-
012	4	6	-
013	4	6	-
019	4	6	-
021	4	6	-
022	4	6	-
023	4	6	-
024	4	6	-
029	4	6	-
031	4	6	-
032	4	6	-
039	4	6	-
040	4	6	-
041	4	6	-
042	-	6	6
043	-	6	6
044	-	6	-
045	-	6	6
046	4	6	-
047	4	6	-
048	4	6	6
049	4	6	6
050	4	6	-
051	-	6	-
052	4	6	-
053	-	6	-
054	4	6	-
055	4	6	-
056	4	6	-
057	4	6	-
058	4	6	-
059	4	6	-
061	-	6	-
068	-	6	-
069	-	6	-
071	-	6	-
072	-	6	-
073	4	6	-
074	-	6	-
078	-	6	-
079	4	6	-

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
081	4	6	-
082	4	6	-
083	4	6	-
084	4	6	-
085	4	6	-
086	4	6	-
087	4	6	-
088	4	6	6
089	4	6	6
091	4	6	-
092	4	6	-
093	4	6	-
094	4	6	-
095	4	6	-
098	4	6	-
099	4	6	-
101	-	6	-
111	3	5	-
118	3	5	-
119	3	5	-
201	4	6	-
203	-	6	-
204	-	6	-
208	-	6	-
209	-	6	-
290	4	6	-
291	-	6	-
292	-	6	-
293	4	6	-
294	4	6	-
295	4	6	-
296	-	6	-
297	-	6	-
298	4	6	-
299	4	6	-
301	3	-	-
302	3	-	-
309	3	-	-
311	3	6	-
312	3	6	-
313	3	6	-
318	3	6	-
319	3	6	-
321	3	6	-
331	3	6	-
332	-	6	-
333	-	6	6
334	3	6	-
338	3	6	6
339	3	6	6

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
401	1	-	-
402	-	6	-
403	-	6	-
404	-	6	-
405	1	-	-
406	1	-	-
407	1	-	-
409	1	6	-
411	-	2	2
412	-	2	2
413	1	2	2
414	-	2	2
415	-	-	2
418	1	2	2
419	1	2	2
421	1	2	2
431	1	-	2
432	1	2	2
439	1	-	2
441	1	2	2
501	3	-	7
502	3	-	7
503	3	-	7
504	3	-	7
509	3	-	7
601	3	6	-
602	-	6	-
603	-	6	6
609	-	6	6
611	-	7	7
621	3	6	-
631	-	7	7
632	-	6	6
633	3	-	7
634	3	-	7
635	3	7	7
636	-	-	6
639	3	7	7
641	-	6	-
642	-	6	-
643	-	6	-
644	3	6	-
649	-	6	-
651	-	6	-
652	-	6	-
653	-	6	-
654	-	6	-
655	-	6	-
659	-	6	-
661	-	7	7

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
662	-	7	7
669	-	7	7
671	-	7	7
678	-	7	7
699	3	7	7
701	3	7	7
711	3	7	7
712	3	-	7
713	3	-	7
714	3	-	7
715	3	7	7
716	3	7	7
718	3	7	7
719	3	7	7
721	3	7	7
722	-	7	7
726	3	-	7
729	3	-	7
731	3	-	7
732	3	-	7
733	3	-	7
735	3	-	7
736	3	-	7
737	3	-	7
738	3	-	7
739	3	-	7
741	3	7	7
742	3	-	7
743	3	-	7
744	3	7	-
745	3	7	7
749	3	7	7
750	3	7	7
751	3	-	7
752	3	7	7
753	3	-	7
754	3	-	7
755	3	-	7
756	3	-	7
757	3	-	7
758	3	7	7
759	3	-	7
761	3	7	7
764	3	7	7
766	-	-	7
767	3	-	7
768	3	7	7
769	3	7	7
771	3	-	7
772	3	-	7

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
774	3	-	7
778	3	-	7
779	3	-	7
781	3	-	7
782	3	-	7
789	3	-	7
791	3	-	7
792	3	-	7
793	3	-	7
794	3	-	7
795	3	-	7
798	3	7	7
799	3	-	7
801	3	7	7
806	3	-	7
808	3	7	7
809	3	7	7
811	3	-	7
812	3	-	7
813	3	-	7
814	3	-	7
818	3	-	7
819	3	-	7
821	3	-	7
822	3	-	7
823	3	-	7
824	3	-	7
825	3	-	7
826	3	7	7
827	3	7	7
828	3	-	7
829	3	7	7
831	3	-	7
834	3	-	7
836	3	-	7
838	3	-	7
839	3	-	7
841	-	-	7
850	3	-	7
851	3	-	7
852	3	-	7
853	3	-	7
854	3	7	7
855	3	7	7
856	3	-	7
857	3	-	7
858	3	-	7
861	3	-	7
871	3	7	7
872	3	-	7

Yrkeskod	Företagare	Tjänstemän	Arbetare
	Socio-ekonomisk grupp		
873	3	-	7
874	3	-	7
875	3	-	7
876	3	-	7
879	3	-	7
881	3	7	7
882	-	7	7
883	-	-	7
888	3	-	7
889	3	-	7
899	3	-	7
901	-	6	6
902	3	6	6
903	-	6	6
904	-	6	6
908	7	7	7
909	7	7	7
911	3	6	-
912	3	8	8
913	-	-	8
914	3	8	8
915	-	8	-
916	-	8	-
917	3	8	-
918	-	8	8
919	3	8	8
921	-	8	8
931	-	8	8
932	3	-	8
933	3	8	8
939	3	-	8
941	3	7	7
942	3	7	7
943	3	7	7
944	3	-	7
945	3	6	-
946	3	6	-
947	3	6	6
948	3	8	8
949	3	8	8
981	-	9	9
999	10	10	10