Published in final edited form as:

J Health Econ. 2015 July ; 42: 29–43. doi:10.1016/j.jhealeco.2015.03.003.

Education and Health: the Role of Cognitive Ability*

Govert Bijwaard[†], Hans van Kippersluis[‡], and Justus Veenman[§]

Govert Bijwaard: bijwaard@nidi.nl; Hans van Kippersluis: hvankippersluis@ese.eur.nl; Justus Veenman: veenman@ese.eur.nl

[†]Netherlands Interdisciplinary Demographic Institute (NIDI-KNAW/University of Groningen), PO Box 11650, 2502 AR The Hague, the Netherlands and IZA, Bonn

[‡]Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands, Tinbergen Institute, Rotterdam, The Netherlands, and Netspar, Tilburg, The Netherlands, +31 10 4088837

[§]Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands, and ERCOMER, Utrecht/Rotterdam

Abstract

We aim to disentangle the relative impact of (i) cognitive ability, and (ii) education on health and mortality using a structural equation model suggested by Conti et al. (2010). We extend their model by allowing for a duration dependent variable (mortality), and an ordinal educational variable. Data come from a Dutch cohort born between 1937 and 1941, including detailed measures of cognitive ability and family background in the final grade of primary school. The data are linked to the mortality register 1995–2011, such that we observe mortality between ages 55 and 75. The results suggest that at least half of the unconditional survival differences between educational groups are due to a ‘selection effect’, primarily on the basis of cognitive ability. Conditional survival differences across those having finished just primary school and those entering secondary education are still substantial, and amount to a 4 years gain in life expectancy, on average.

Keywords

Education; Cognitive Ability; Mortality; Structural Equation Model; Duration model

*Van Kippersluis gratefully acknowledges funding from the National Institute on Aging (NIA) under grant R01AG037398, from NETSPAR under the project “Income and health, work and care across the life cycle II”, and from the Netherlands Organization of Scientific Research (NWO Veni grant 016.145.082). The authors acknowledge access to linked data resources (DO 1995–2011) by Statistics Netherlands (CBS). We thank Mars Cramer and Mirjam van Praag for help in accessing the original data (see <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:39042>), and are grateful to Mars Cramer, and attendants from the Empirical Health Economics conference in Munich 2013, the International Health Economics Association conference in Sydney 2013, the Multistate Event History Analysis in Hangzhou 2012, and seminar participants at the Chinese University of Hong Kong, Cornell University, Erasmus University Rotterdam, and the University of Southern California for helpful comments.

© 2015 Published by Elsevier B.V.

Correspondence to: Hans van Kippersluis, hvankippersluis@ese.eur.nl.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 Introduction

Disparities in health and life expectancy across educational groups are striking and pervasive, and are considered one of the most compelling and well established facts in social science research (Mazumder, 2012). Even in an egalitarian country such as the Netherlands, with a very accessible health care system, the difference in life expectancy between the university educated and those who finished only primary school is 6 to 7 years (CBS, 2008). It is commonly assumed that a large part of this association derives from the causal effect of education on health outcomes. An abundant list of possible mechanisms was proposed, among which occupational demands, health behavior, and the ability to process information are the most commonly mentioned (Ross and Wu, 1995; Cutler and Lleras-Muney, 2008).

Yet, the association between education and health could also stem from (i) ‘reverse causality’, in which childhood ill-health constrains educational attainment (Behrman and Rosenzweig, 2004; Case et al. 2005), and (ii) confounding ‘third factors’ such as ability, parental background and time preference that influence both education and health outcomes (Fuchs, 1982; Auld and Sidhu, 2005; Deary, 2008).

Studies based on natural experiments in education, such as changes in compulsory schooling laws, overcome the difficulty of separating the direct causal effect of education from third factor effects. The estimates based on these studies point towards a small effect (Lleras-Muney, 2005; Oreopoulos, 2006; Van Kippersluis et al. 2011; Meghir et al. 2013), or even insignificant effect of education on health and mortality (Arendt, 2005; Albouy and Lequien, 2008; Mazumder, 2008; Braakmann, 2011; Clark and Royer, 2013). This suggests that confounding factors may well play an important role in shaping the strong association between education and health.

Surprisingly little research in economics has investigated the contribution of early childhood abilities and childhood social background in shaping the association between education and health.¹ Some recent economic studies report associations between childhood cognitive and non-cognitive abilities, and health outcomes at ages 30–40 using the British Cohort Study (Murasko, 2007), the U.K. National Child Development Study (Carneiro et al. 2007), the U.S. National Longitudinal Study of Youth 1979 (Auld and Sidhu, 2005; Kaestner and Collison, 2011), or the Dutch ‘Brabant data’ (Cramer, 2012). It is established that cognitive ability and some non-cognitive factors such as self-esteem and conscientiousness are associated with health outcomes. Nonetheless, hardly anything is known about (i) the relative impact of education and childhood abilities on health outcomes, and in turn (ii) how much of the association between education and health is explained by these cognitive and non-cognitive abilities.

A notable contribution to the literature is a recent series of papers by Conti and Heckman (2010), Conti et al. (2010; 2011), and Heckman et al. (2014) who, using the British Cohort Study and the National Longitudinal Study of Youth (NLSY79), estimate a structural equation model in which the interdependence between education, health, and two latent

¹See Gottfredson (2004) for an overview of the epidemiological literature.

factors capturing cognitive and non-cognitive abilities is explicitly modeled. The authors show that for most health outcomes around half of the association between education and health is driven by cognitive and non-cognitive abilities and early childhood social background. The other half is interpreted as the causal effect of education on health.

While the series of papers by Conti, Heckman and co-authors provided a significant contribution to the literature, there are two notable limitations. First, the health outcomes are measured at age 30, an age at which health differences by education may not have fully materialized. In fact, disparities in health and mortality seem to peak around middle-age (Cutler and Lleras-Muney, 2008). Secondly, the health measures are all self-reported, which may bias the estimates since education is related to subjective health perceptions (Bago d'Uva et al. 2008).

In this paper we aim to disentangle the effects of education and cognitive ability on health outcomes. We will use the so-called 'Brabant data' - a representative cohort of primary school sixth graders in the Dutch province of Noord-Brabant - that has detailed information on cognitive ability and social background measured back in 1952. Three follow-up surveys in 1957, 1983 and 1993 contain information on education, employment, and self-reported health. We have linked these data to the mortality register 1995–2011, such that the impact on mortality can be analyzed.

The contribution of this paper is threefold. First, we study the relative impact of cognitive ability and education on mortality, as an *objective* health indicator. The second contribution is that, in contrast to existing studies that measure health outcomes at ages 30–40, we observe mortality during ages 55–75. Finally, we extend the structural equation model by Conti et al. (2010) by allowing for a duration dependent variable (mortality).²

The results show that for most ages, cognitive ability and family socioeconomic status explain around half of the raw differences in mortality across educational groups. Stated otherwise, education remains important in determining mortality even after controlling for cognitive ability, family socioeconomic status, and a range of other background variables. The conditional survival differences across educational groups are even remarkable, and amount to a 4-year gain in life expectancy for those entering at least secondary school compared to those that dropped out after primary school.

This paper is structured as follows. Section 2 presents the Brabant data including the available register data from Statistics Netherlands, section 3 presents the structural equation model that we will use to disentangle the relative contributions of cognitive ability and education on health outcomes. Section 4 presents the results and section 5 discusses them.

²Savelyev (2012) developed a similar structural equation model for mortality as ours, yet using a discrete-time hazard model and not taking into account dynamic selection of ability due to differential survivorship. His study is based on the Terman data, a cohort of individuals with IQ beyond 140. Hence, apart from differences in the model specification, his focus is on an extraordinary sample corresponding to the 99.6th percentile of the intelligence distribution, with very limited variation in cognitive ability. Not surprisingly, he examines the effect of higher education whereas we focus on secondary education.

2 Data and descriptive statistics

The data are from a Dutch cohort born between 1937 and 1941. Very detailed information about individual intelligence, social background and school achievement is available for 5,823 individuals. The survey was held in the spring and summer of 1952 among pupils of the sixth (last) grade of primary schools in the Dutch province of Noord-Brabant, and hence is referred to as the ‘Brabant data’. One-fourth of the province population was sampled; mainly by including every fourth child from the schools’ list of pupils.³ Hartog (1989) investigated the data and found no reason to doubt representativeness. A selective dropout of pupils before participating in the data collection does not exist, as primary school was compulsory and enforcement of school attendance was strict (Dronkers, 2002).

Follow-up surveys took place in 1957, 1983 and 1993.⁴ In 1957 only a sub-sample - those who scored above-average on six tests - of the original cohort was interviewed about the school careers between 1952 and 1957 to particularly investigate school career choices of the most intelligent half of the cohort. In 1983 and 1993 attempts were made to trace all initial respondents of the Brabant-cohort to investigate labour market behavior, with overall response rates of around 45 percent. The sample is reduced to 2,998 individuals who have measurements in 1952 and in either 1983 or 1993, or both.⁵

The Brabant data are subsequently linked to administrative records from Statistics Netherlands. The basis for this linkage is identifying information on ZIP code, date of birth, and sex, provided in 1993 by Dutch municipalities, which includes information on all individuals living in the Netherlands. The administrative records are available since 1995. Because of the two-year discrepancy only 86 percent of the 2,998 individuals could be traced in the municipality register in 1995, leaving us with a working sample of 2,579 individuals. Administrative records include the mortality register and the municipality register for the years 1995–2011 inclusive. The mortality register is used to identify drop out due to death in the follow-up period. Demographics are obtained from the municipality register.

Dependent variables

Our outcome variable is *Mortality*, which is identified from the mortality register in the period 1995–2011. Given that most pupils are born around 1940, this implies that we follow mortality from age 55 until 75.⁶ In our sample, 409 individuals, or 16 percent, died during the period 1995–2011. Close to 50 percent died from cancer, 25 percent from cardiovascular diseases, and 8 percent from respiratory diseases such as COPD and pneumonia. External causes such as accidents comprise only two percent, as do mental disorders (e.g. dementia),

³Some schools had school years beginning in April rather than in September. For these schools, half the pupils of half the schools were included in the sample, which yielded 369 observations on a total of 5,823 (Hartog, 1989).

⁴Mathijssen and Sonnemans (1958), Hartog and Pfann (1985), Van Praag (1992), and Hartog, Jonker and Pfann (2002). The complete questionnaire is included in Van Praag (1992) ‘Brabantse zesdeklassers, 1952–2010’.

⁵In section 4.2 it is verified that selective attrition does not affect our results.

⁶Of the Dutch population 1940 cohort, only 6.8 percent died between the ages of 12 and 55 – Human Mortality Database, University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on July 30, 2012).

diseases of the digestive system (e.g. liver cirrhosis) and diseases of the nervous system (e.g. Parkinson).

Independent variables

Our main independent variable of interest is *Education*, here defined as the highest level of education attended, in three categories: (1) *Lower Education*, including those who attended at most (extended)⁷ primary school, (2) *Lower Vocational Education*, including those who attended at most lower vocational education such as the lower agricultural school or lower polytechnic schools, and (3) *At least General Secondary School*, including those who attended lower general secondary school, higher general secondary school, and higher vocational education or university. Education is retrieved mainly from the 1983 and 1993-survey variables on the highest level of education attended. The maximum of the two defines *Education*, and where missing we update our educational variable with information from the 1957 survey.

Table 1 presents descriptive statistics and shows that 14 percent did not continue school after primary school forming the *Lower Education* category, 35 percent only attended *Lower Vocational Education*, and the other 51 percent attended *At least General Secondary School*. Figure 1 shows the Kaplan-Meier survival curves for a binary indicator of education with threshold at *Lower Education*, and separately for the three education categories. It is clear that the largest survival differences are between those with only primary school and those above primary school, and that the difference grows with age to around ten percentage points near age 75.

Our second independent variable is *Cognitive Ability*. In the Brabant data there are two measurements for cognitive ability, both measured in the final grade of primary school (i.e. around age 12): (i) the Raven Progressive Matrices Test, and (ii) a Vocabulary test (picking synonyms).⁸ The timing of the intelligence tests implies that the plausible feedback effects from education to cognitive ability (Deary and Johnson, 2010; Brinch and Galloway, 2012; Meghir et al., 2013) will be seen as an education effect, while the impact of the cognitive ability endowment in the final grade of primary school on educational choice and later-life mortality is seen as a selection effect.⁹

The IQ p.m. ('progressive matrices') test focuses on mathematical ability and is a replication of the British Progressive Matrices test, designed by Raven (1958). It is considered to be a 'pure' measurement of problem solving abilities, as it does not require any linguistic or general knowledge (Dronkers, 2002). Hence, the Raven test is supposed to measure fluid or analytic intelligence (Carpenter et al. 1990). In this sense, the test can be compared to

⁷At the time, pupils had to stay in school for at least 8 years, or until they reached the age of 14. Since regular primary school only consisted of 6 grades, some schools offered an additional 2-year extended primary school ("vglo").

⁸The data also contain the so-called LO-IV test, which consists of six sub-tests: regularities in series of numbers, analogies in figures, analogies in words, and similarities between concepts (equal, not-equal, cause). Since the quality of this test has been questioned (Hartog et al. 2002, p. 5) we will not use it in our analyses. There is also information on grades for specific courses (Dutch language, mathematics (arithmetics), history, physics, geography, health sciences, and traffic), but since these are not clean measures of cognitive ability and are relative to others in one's classroom, we choose not to use these grades.

⁹It should be emphasized however that there could be unobserved factors correlated to both cognitive ability in the final grade of primary school and later-life mortality, which our measure of cognitive ability would be picking up.

Spearman's g test (1927). The term g refers to the determinants of the common variance within intelligence tests, being the core issue of intelligence measurement (Carpenter et al. 1990).

Table 1 shows that the ability test designed by Raven has an average of 102, with standard deviation of 13 while the vocabulary test is 101, on average, with standard deviation 13. The correlation between the Raven test and the vocabulary test is 0.38. This suggests that while there seems to be some overlap between the two measurements, the tests additionally gauge some idiosyncratic part of cognitive ability. Therefore, we will use both measurements to build a comprehensive latent factor of cognitive ability. In a robustness check we solely use the Raven test to see whether the results differ.

Control variables

Apart from a fairly standard set of demographic control variables such as *Age*, whether *Male*, and *Birth Rank*, we also have information about the social and school environment of the individuals. Most of these variables are reported by the School Principal. *Family Socioeconomic Status* is measured in three categories from lowest to highest depending on father's occupation.¹⁰ We additionally know whether the child had to work in the parent's farm or company, defining the binary indicator *Child Works*, which potentially signals part of the childhood health status. In this (historical) case, however, the variable is mainly dependent on the parents having a firm.

Available information regarding the school includes *School Type* and the *Number of Teachers*. *Repeat* defines the number of classes that children had to repeat. Further, we know the *Teacher's Advice* regarding further education of the child, and the *Preference of the Parents* concerning the education of the pupil, categories of which are defined in Table 1, which also includes descriptive statistics.

We have no information about childhood health status, which prevents us from investigating the possibility of reverse causality from health to education in our sample. The sample is comprised of pupils who made it to the final grade of primary school. Hence, pupils with severe health problems impairing going to school in the first place will not be present in our sample. Moreover, in the 1983 wave of the survey male respondents were asked whether they served in the military. The main reason for disqualification of compulsory military duty is health problems.¹¹ Since the fraction of individuals having served in the military is almost identical across educational levels, this provides some indirect evidence that health differences across educational levels were minimal during teenage years. We furthermore refer to Conti et al. (2010) who showed that in their sample childhood health, as measured

¹⁰We classify lower administrative, agricultural, industrial, and other lower workers, and the disabled into the Lowest Socioeconomic Status. If the School Principal considered the family antisocial, the family is also classified into the Lowest Socioeconomic Status. Intermediary personnel, self-employed farmers, self-employed craftsmen, and the retired are categorized into the Intermediate Socioeconomic Status (following Cramer, 2012). Teachers, executives and academics are classified into the Highest Socioeconomic Status. In case father's occupation is missing, we use father's education for individuals in the 1957 survey. Father's education is classified into 3 levels, which we directly translate into the three socioeconomic statuses. We use mother's education in case the father died or was not present in the household.

¹¹Other reasons were exemption owing to one's brother's service, grounds of conscience, or personal indispensability (e.g. Van Schellen and Nieuwbeerta, 2007).

by childhood height, was not an important determinant of educational choice. The lack of information on childhood health should therefore not be a major source of concern.

3 Methodology

Our empirical approach is an extension of the structural equation framework developed by Conti et al. (2010). We briefly describe the Conti et al. model, after which we will present our two extensions: allowing for an objective duration dependent variable (mortality), and introducing an ordinal educational choice. Finally, we explain how we disentangle the effects of cognitive ability and education on the health outcomes.

3.1 Basic structural equation model

The basic Conti et al. model allows a way of modeling the interrelationships between abilities, education and health outcomes, where individuals potentially make their educational decisions depending on the perceived health gains. Hence, the educational choice is endogenous, and in practice it is assumed that selection into schooling can be fully accounted for by using observed characteristics and unobserved ability. The model consists of three parts: (i) a binary educational choice depending on latent abilities and other covariates, (ii) potential outcomes depending on the choice of education, latent abilities, and other covariates, and (iii) a measurement system for the latent abilities.

The binary indicator for education D_i is defined as 1 if individual i took any education beyond the compulsory schooling age, and 0 if not:

$$D_i = \begin{cases} 1 & \text{if } D_i^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where we assume D_i^* is an underlying latent utility which is continuous and linear, and depends on latent abilities θ , and observed characteristics X^D :

$$D_i^* = \gamma X_i^D + \alpha_D \theta_i + v_{iD} \quad (2)$$

with v_{iD} being an error term independent of X^D and θ . We assume that v_{iD} is normally distributed, which implies that we have a probit model for the educational choice. We fix the variance at 1 since the variance is not identified in a probit model.

The second part is the potential outcomes part, in which there are two potential outcomes Y_{i1} and Y_{i0} , where the former is the outcome in case the individual chose to pursue education beyond what is compulsory, and the latter is the outcome in case the individual dropped out of school right after the compulsory schooling age. Both Y_{i1} and Y_{i0} depend on latent ability θ , and on observed characteristics X^Y :

$$Y_{i1} = \beta_1 X_i^Y + \alpha_1 \theta_i + v_{i1} \quad (3)$$

$$Y_{i0} = \beta_0 X_i^Y + \alpha_0 \theta_i + v_{i0} \quad (4)$$

with (v_0, v_1) independent of X^Y and θ , independent of v_{iD} conditional on X^Y and θ , and both follow a normal distribution with variance σ_1^2 and σ_0^2 , respectively.

The final part of the model is the measurement equation, where one or two measurements, M_{ik} ($k = 1, 2$), implicitly define the latent ability θ :

$$M_{ik} = \delta_k X_i^M + \alpha_{M_k} \theta_i + v_{iM_k} \quad (5)$$

with v_{M_k} independent of X^M and θ . We assume that v_{M_k} is normally distributed with variance $\sigma_{M_k}^2$.

3.2 Allowing for a duration outcome as dependent variable

While the basic model is useful in disentangling the relative contributions of education and abilities on continuous and binary health outcomes, it does not allow for a duration outcome like survival till death.

In our extended model, the first part is the same, defining a binary educational choice as in (1) and (2), placing the cut-off at *Lower Education* (primary school). Hence, in our model individuals face the choice of quitting after primary education ($D = 0$), or enrolling into secondary education ($D = 1$). The measurement equation for latent ability is defined by (5), where we have two measurements for latent cognitive ability.

It is common practice to define the potential outcomes of a duration variable like mortality in terms of the hazard that the outcome of interest occurs.¹² We define $\lambda^{(1)}(t)$ as the hazard rate for an individual with education level beyond primary school ($D_i = 1$), and $\lambda^{(0)}(t)$ as the hazard rate for an individual with an education level equal to primary school ($D_i = 0$). We assume a Gompertz proportional hazard model for the two potential hazards, which has been shown to be an accurate representation of mortality between the ages of 30 and 80 (e.g. Gavrilov and Gavrilova, 1991; Cramer, 2012). Both potential hazards depend on the latent ability θ ,¹³ and observed characteristics X^Y :

$$\lambda^{(0)}(t|X^Y, \theta) = \exp(a_0 t + \beta_0 X_i^Y + \alpha_0 \theta_i) \quad (6)$$

$$\lambda^{(1)}(t|X^Y, \theta) = \exp(a_1 t + \beta_1 X_i^Y + \alpha_1 \theta_i) \quad (7)$$

The effect of latent ability on the hazard is captured by α_0 and α_1 . The corresponding potential survival rates are

¹²We can use a duration model with potential outcomes because the endogenous education choice is determined before mortality plays a major role: mortality can be largely ignored for young ages. If the education choice would still play a role during higher mortality rates the model for educational choice should take selective survival effects into account. Then a ‘timing-of-events’ model could be a better model, see Abbring and Van den Berg (2003).

¹³The latent ability in the hazard is similar to including unobserved heterogeneity in the hazard, and for identification the unobserved heterogeneity needs to have a finite mean. The mean of the unobserved heterogeneity term in our model, $e^{a\theta}$, only depends on a and is finite when a is finite.

$$S^{(0)}(t|X^Y, \theta) = \exp \left(- \int_0^t \lambda^{(0)}(s|X^Y, \theta) ds \right) \quad (8)$$

$$S^{(1)}(t|X^Y, \theta) = \exp \left(- \int_0^t \lambda^{(1)}(s|X^Y, \theta) ds \right) \quad (9)$$

Without additional restrictions on the distribution of the latent factors the model is not identified. However, because we have an intrinsically non-linear duration outcome instead of a linear outcome, the Ledermann bound on the number of measurements compared to the number of latent factors does not apply. Identification of our model is closely related to the identification in a mixed proportional hazard (MPH) model, where we assume that the unobserved heterogeneity has a log-normal distribution. A MPH model is identified when the unobserved heterogeneity term has finite mean and is independent of the other observed factors (Elbers and Ridder, 1982). When we assume a normal distribution for the latent ability, $\theta \sim \mathcal{N}(0, \sigma_\theta^2)$, the implied unobserved heterogeneity in the hazard (equations (6) and (7)) has mean $\exp(\frac{1}{2}\alpha_j^2 \sigma_\theta^2)$ for $j = 0, 1$. For identification α_j or σ_θ^2 needs to be fixed. We choose to fix $\sigma_\theta^2 = 1$. Thus the latent ability follows a standard normal distribution.¹⁴

An important feature of duration data is that for some individuals we only know that he or she survived up to a certain time (often the end of the observation window). In this case an individual is (right) censored, $\Lambda_i = 0$, and we use the survival function instead of the hazard in the likelihood function. Another feature of duration data is that only individuals are observed having survived up to a certain age. In our case, mortality follow-up is only available from age 55 onwards. In this case the individuals are left-truncated, and we need to condition on survival up to the age of first observation, t_0 .

The likelihood contribution of individual i in our duration model is

$$L_i^{(j)} = \lambda^{(j)}(t)^{\Lambda_i} S^{(j)}(t) / S^{(j)}(t_0), \quad j=0, 1 \quad (10)$$

With left-truncated data the distribution of latent ability among the survivors (up to the left-truncation time) changes. When only individuals are observed that have survived until age t_0 the likelihood contribution is

$$\begin{aligned} L_i = & \int \left[\Phi(\gamma X_i^D + \alpha_D \theta) \cdot \lambda^{(1)}(t|X^Y, \theta)^{\Lambda_i} S^{(1)}(t|X^Y, \theta) / S^{(1)}(t_0|X^Y, \theta) \right]^{D_i} \\ & \times \left[\Phi(-\gamma X_i^D - \alpha_D \theta) \cdot \lambda^{(0)}(t|X^Y, \theta)^{\Lambda_i} S^{(0)}(t|X^Y, \theta) / S^{(0)}(t_0|X^Y, \theta) \right]^{1-D_i} \\ & \times \prod_{k=1}^2 \frac{1}{\sigma_{M_k}} \phi \left(\frac{M_{ik} - \delta_k X_i^M - \alpha_{M_k} \theta}{\sigma_{M_k}} \right) dH(\theta | T > t_0) \end{aligned} \quad (11)$$

¹⁴In principle restricting the distribution of θ to a normal distribution is not necessary. In line with the literature on MPH models a discrete distribution with finite points of support would be an alternative choice (Heckman and Singer, 1984). However, using a normal distribution assumes a continuum of ability values rather than a finite number, and the distribution of intelligence is generally found to be close to normal (Gottfredson, 1997).

with the distribution of the latent abilities conditional on survival up to t_0

$$dH(\theta|T>t_0) = \frac{\left[\Phi(\gamma X_i^D + \alpha_D \theta) S^{(1)}(t_0|X^Y, \theta) + \Phi(-\gamma X_i^D - \alpha_D \theta) S^{(0)}(t_0|X^Y, \theta) \right] h(\theta)}{\int \left[\Phi(\gamma X_i^D + \alpha_D \theta) S^{(1)}(t_0|X^Y, \theta) + \Phi(-\gamma X_i^D - \alpha_D \theta) S^{(0)}(t_0|X^Y, \theta) \right] h(\theta) d\theta} \quad (12)$$

with $h(\theta)$ is a normal distribution with variance $\sigma_\theta^2=1$. The maximum likelihood estimation of the parameters involves the calculation of an integral that does not have an analytical solution. However, Gaussian quadrature can approximate this one dimensional integral very well. Hence, we estimate the parameters using maximum likelihood on the basis of Gaussian quadrature approximation.¹⁵

3.3 Allowing for an ordered discrete educational choice

Usually education is available in more than two categories with a natural ordering of the alternative education levels. As a robustness check (see section 4.2), we extend the standard model to account for this type of ordinal independent variable, where the starting point is, again, an index model with a single latent variable given as in (2). Assume there are K education levels and define D_i as the indicator of education that takes value k if the individual has reached education level k :

$$D_i=k \quad \text{if } \zeta_{k-1} < D_i^* \leq \zeta_k \quad (13)$$

where $\zeta_0 = -\infty$ and $\zeta_K = \infty$. Then, assuming normally distributed v_D , we have an ordered probit model with $(K-1)$ additional threshold parameters, ζ_k . Each education level now has a corresponding potential Gompertz hazard $\lambda^{(k)}$, that depends on exogenous characteristics X^Y and on the unobserved latent ability, θ , i.e.,

$$\lambda^{(k)}(t|X^Y, \theta) = \exp(a_k t + \beta_k X_i^Y + \alpha_k \theta_i) \quad (14)$$

3.4 Disentangling the effects of ability and education

At the individual level, the main estimand of interest is the survival difference across the two educational levels, $S^{(1)}(t) - S^{(0)}(t)$, where $S^{(1)}(t)$ denotes the survival time up to age t for individuals with at least secondary education ($D=1$), and $S^{(0)}(t)$ is the survival time up to age t for those with primary school only ($D=0$). We are interested in the expected value of this identity for a given (sub) population. In the sample, the difference in the Kaplan-Meier survival curves is the *unconditional* survival difference between the two levels of educational attainment, $E[S^{(1)}(t) - S^{(0)}(t)]$. This unconditional difference can be interpreted as the association between education and mortality.

¹⁵Gaussian quadrature is a numerical integration method based on Hermite polynomials (Press et al. 1993). It provides an efficient approximation for evaluating indefinite integrals based on normal distributions (Butler and Moffitt, 1982). A similar method has been applied before in survival analysis (Lillard, 1993).

Here we are interested to what extent this association is driven by cognitive ability and other control variables. Using the estimated parameters, we define the *conditional* survival difference between the two levels of educational attainment, where conditioning is based on cognitive ability and the other control variables, as follows:

$$\int \int E \left[S^{(1)}(t) - S^{(0)}(t) | X=x, \theta=c \right] dF_{X,\theta}(x, c) \quad (15)$$

where X are the covariates, and θ is the value of latent cognitive ability. We integrate over the joint distribution of the covariates and latent ability, $F_{X,\theta}(x, c)$.¹⁶ Note that these conditional survival differences are conditional on surviving to the initial age, which is 55 in our case.

Unfortunately, the integrals cannot be solved analytically, as the dimension of the covariates X is too large. Another issue is that the comparison of the survival functions involves the counterfactual of surviving with another education level. Hence in order to illustrate the conditional survival differences we resort to simulation.¹⁷ For each education level we simulate the survival of 10,000 individuals. To each individual we assign observed characteristics based on the empirical distribution in the sample. The simulation procedure consists of four steps:

1. Draw a vector of parameter estimates assuming that the estimator is normally distributed around the point estimates with a variance-covariance matrix equal to the estimated one.
2. Compute the conditional hazard rates based on these parameter values and individual characteristics using (6) and (7), conditional on the value of the latent ability.
3. Determine the unconditional survival function for every individual and for the whole age-range from 55 to 100 on the basis of equations (8) and (9), and by integrating out the latent ability through Gaussian quadrature methods.
4. Calculate the average (over the 10,000 individuals) survival at each age (with steps of a month).

We repeat these steps 100 times to obtain 100 independent observations of the survival function for each education level.

With this information, we can compute the fraction of individuals who are still alive at a certain age for the two educational groups (both the average and the variance). This defines the conditional survival difference between the two educational groups, since we condition

¹⁶Since the conditional survival differences may well be very different for individuals in different parts of the education distribution, we additionally define the conditional survival difference for those with $D = j, j = 0, 1$ as follows:

$$\int \int E \left[S^{(1)}(t) - S^{(0)}(t) | X=x, \theta=c, D=j \right] dF_{X,\theta|D=j}(x, c)$$

¹⁷Cockx and Picchio (2012) use a similar simulation procedure to obtain predictions and the standard errors for non-linear combinations of the parameters. Elbers et al. (2003) used a related method to account for counterfactual uncertainty.

on cognitive ability and the other covariates. The simulations also allow us to compute life expectancy (and its standard error) separately for the two educational groups, by multiplying the survival function with the age steps.

In order to illustrate the relative importance of education and cognitive ability, we decompose the unconditional survival differences from the Kaplan-Meier curves in Figure 1 into the conditional survival difference and a residual, which is a selection effect on the basis of cognitive ability and the other observable factors. Mathematically,

$$E \left[S^{(1)}(t) - S^{(0)}(t) \right] = \int \int E \left[S^{(1)}(t) - S^{(0)}(t) | X=x, \theta=c \right] dF_{X,\theta}(x, c) + \varepsilon(X, \theta) \quad (16)$$

where the LHS represents the unconditional survival differences represented by the Kaplan-Meier survival curve, the first part of the RHS is the conditional survival difference defined in (15), and $\varepsilon(X, \theta)$ represents the selection effect on the basis of observable characteristics X and cognitive ability θ . Note that this selection effect is the combination of actual selection bias and selection based on perceived gains of secondary education.

For the ordinal education measure the procedure is very similar. We have three potential hazards and three possible survival functions, one corresponding to each educational level. Although there are more possibilities now to compare the educational groups, we choose to focus on two binary comparisons of the particular educational level to the educational level directly preceding it. Hence, we estimate two different conditional survival differences: (i) lower vocational education compared to primary education only, and (ii) at least general secondary education compared to lower vocational education.

4 Results

Our baseline specification is the survival model with a binary education variable and two measurements for cognitive ability. We estimate the model by maximizing the likelihood in (11), and present the results in section 4.1. Exogenous factors influencing the outcome, X^Y in (6) and (7), include male, whether the child is working, family socioeconomic status, and birth rank. Factors *additionally* influencing the measurements of cognitive ability, X^M in (5), include school type and the number of teachers at school. Finally, on top of the exogenous variables affecting the outcome and intelligence, *additional* factors influencing the educational choice, X^D in (2), include the teacher's advice, whether a grade was repeated, and the preference of the parents.

4.1 Main results

Table 2 contains the parameter estimates of the model. The first column shows that our latent factor of cognitive ability strongly influences the educational choice, as expected. The probability of entering secondary school can be derived from the impact of the latent factor and is already beyond 0.6 for those with the lowest cognitive abilities, and gradually increases towards 1.0 for those with the highest cognitive abilities (see Figure 2).

Conditional on other observed characteristics such as parental preference, teachers advice, and family socioeconomic status, males were less likely to enter secondary school, as are children who had to work in the family business during primary school. Family socioeconomic status is a strong predictor of education, with children from families with a higher socioeconomic status significantly more likely to enter secondary school. Children who went to protestant or other schools, as compared to those who went to catholic schools, were more likely to enter secondary school. Strong predictors of educational choice are the teacher's advice and the preference of the parents. Children who repeated one or more grades were less likely to enter secondary school.

Columns 2 and 3 show that on both measurements of cognitive ability girls did slightly better, and children from families with a higher socioeconomic status had higher scores. School characteristics such as the school type and the number of teachers also relate to the test scores.

The final two columns of the table present the determinants of mortality across the two educational groups. While the point estimates of the effect of cognitive ability on mortality are negative as expected, the effects do not reach statistical significance at the 10 percent level, although the p-values are close to the 10 percent cut-off. The point estimates suggest that the effect of a one-standard deviation increase in cognitive ability is to reduce the mortality hazard by 18% and 28%, for those with primary school only and those with at least secondary education, respectively. These results are comparable to both the results presented in the review by Batty et al. (2007) for the Scottish Mental survey for individuals born in 1921 whose intelligence was measured at age 11, with mortality follow-up for 65 years (reduction in hazard rate 21%), and the results presented by Batty et al. (2009) for a cohort of one million Swedish whose IQ was measured at age 18 and who were followed for 20 years (reduction in hazard rate 25%). Males have a higher hazard of dying compared to females, although the effect is only statistically significant among the higher educated.

The coefficients in Table 2 allow computing the conditional survival difference across educational groups, as described in section 3.4. Figure 3 shows these conditional survival differences for all age groups from 55 to 75 years of age. The survival difference between the two educational groups, conditional on family background and cognitive ability, is positive and increases with age. Note that the confidence intervals are fairly wide, such that the conditional survival differences only reach statistical significance at higher ages. The sizes can be interpreted as percentage point differences in the survival probability at a certain age. Hence, around age 70, entering secondary school is associated with a 2 percentage point increase in the survival probability (90% confidence interval -1 to 7). Around age 74, entering secondary school is associated with a 5 percentage point increase in the survival probability (90% confidence interval 1 to 10).

If we extrapolate the estimated survival functions outside of our observed age window, the simulations allow computing the estimated differences in life expectancy for the average individual in the sample. This provides an alternative summary measure. The life expectancy of those only finishing primary school is 82.86 (standard error 1.37), compared with 87.15 (1.11) for those having finished at least secondary school,¹⁸ a statistically

significant difference of 4.29 years (1.58). This implies that entering secondary school is associated with an increase of more than 4 years in life expectancy, which is within the bandwidth of the raw survival difference of 5 years across individuals with primary and secondary education. It has to be acknowledged, however, that this estimate is based upon extrapolation and hence on relatively strong functional form assumptions.

We decompose the unconditional differences in the Kaplan-Meier survival curves from Figure 1 into a conditional difference and a selection effect based on cognitive ability and other observable control variables.¹⁹ Figure 4 shows that at early ages mortality differentials are mainly due to selection effects, while after age 60 the importance of education increases. For most ages, the selection effect is responsible for around half of the unconditional differences in survival across educational groups. It has to be acknowledged though that these conclusions are based on the point estimates, uncertainty around which is relatively large in this case.

To gauge the importance of cognitive ability in the selection effect, we additionally ran all models without the latent factor for cognitive ability. The results show that the conditional survival differences are larger in a model without cognitive ability.²⁰ This is an indication that cognitive ability plays an important role in the selection effect. It is tempting to decompose the selection effect into a selection due to cognitive ability and a selection on other observable characteristics. The selection on other observable characteristics can be computed as the difference between (i) the unconditional difference from the observed Kaplan-Meier survival rate of the two education levels and (ii) the conditional survival difference from the model without cognitive abilities. The selection on cognitive ability can then be easily computed as the difference between the total selection effect and the part of the selection effect attributed to other observable characteristics. This is illustrated in Figure 5, which shows that cognitive ability explains the largest part of the selection effect. In fact, selection on other observable factors is even negative between ages 60 and 70. We have to emphasize, however, that this interpretation should be taken with care as cognitive ability could be correlated to other control variables in the model, and as such the selection effect may not be additive.

4.2 Robustness checks

In a model with an ordinal educational choice, corresponding to the three levels in the definition of *Education* in section 2, the coefficient estimates of the exogenous variables are very similar to the ones presented for the binary educational variable.²¹ Figure 6 presents the conditional survival differences for the three different educational levels. It is clear that there is a large, but insignificant, conditional survival difference between lower vocational school (level 2) and primary school (level 1). At age 75, those who only attended primary

¹⁸These outcomes are reasonably close to the gender-education specific estimated life expectancies conditional on surviving to age 55 that Statistics Netherlands presents on the basis of the 2012 mortality risks (see <http://www.cbs.nl/en-GB/menu/themas/gezondheid-welzijn/cijfers/extra/resterende-gezonde-levensverwachting.htm?Languageswitch=on>). Men and women having finished primary education can expect to live up to 78.1 and 82.2, respectively. Men and women having finished higher education can expect to live up to 83.3 and 87.1, respectively.

¹⁹The corresponding graphs using the distribution of X those with $D = 1$ and $D = 0$ are very similar.

²⁰Results are available upon request.

²¹All results not presented and the details of the models used in this section are available upon request.

school are around four percentage points more likely to die than those who attended lower vocational school. The conditional survival difference between general secondary school and lower vocational school is practically zero.

If we decompose the unconditional survival differences between the three educational groups into a conditional survival difference and a selection effect, we obtain Figure 7. This graph shows that the conditional survival difference between primary and vocational education is positive and becomes larger than the selection effect from age 70 onwards, in line with the findings of the dichotomous indicator for education. The conditional survival difference between vocational and higher education is negligible. Taken together, Figures 6 and 7 clearly indicate that the largest difference is between those having finished primary school and those beyond primary school, such that the dichotomization in the previous subsection seems justified.

While mortality is an objective, and in some sense ‘the ultimate’, health outcome, the influence of education and cognitive ability may differ depending on the health outcome used. In the 1993 wave of our Brabant survey, hence around age 53 for our sample, a subjective assessment of one’s health was asked to the respondents in five categories, i.e. ‘poor’, ‘sometimes good, sometimes bad’, ‘fair’, ‘good’, and ‘very good’. We estimated the model described in section 3, now allowing for an ordinal dependent variable, to check robustness to our main outcome measure, and to compare our results to the literature.

We estimated the conditional difference in the probability to report any of the five categories between the two educational groups. The conditional difference in the probabilities for the categories ‘poor’ and ‘sometimes good, sometimes bad’ are significantly lower with -0.03 (standard error 0.007) and -0.08 (0.011) respectively. The conditional difference in the probabilities for the categories ‘fair’ and ‘very good’ are very close to zero, while the conditional difference in the probability of reporting to be in ‘good’ health between those having finished only primary school and those entering secondary education is large (and significant) and amounts to a 15 (standard error 0.011) percentage points increase.

When comparing our results to the literature, we confirm the findings of Hartog and Oosterbeek (1998) that both education and cognitive ability affect self-reported health. Conti and Heckman (2010) used a binary indicator for ‘poor health’ and found that half of the raw differences in poor health is due to a treatment effect of education and the other half was selection. We find that selection is not as important as in our mortality analyses, while education does play a large role in explaining the raw differences in health levels in a model with five health levels and binary education, see Figure 8. This suggests that the relative contributions of education and selection effects may well differ across objective health measures and the subjective health measures that are commonly used in the literature.

Since the sample size is somewhat small we chose not to present all results separately by gender. Yet, since both educational choices and survival are obviously dependent on gender, we ran all models separately for males and females. Strong disparities in survival across educational groups exist for both males and females. This can be distracted from Figure 9 where the height of the bar indicates the unconditional survival differences across the two

educational groups. While unconditional survival differences across educational groups are larger for females, the relative importance of education, derived from the decomposition of the raw survival differences is higher for males.

One could argue that the Raven progressive matrices test is a purer measurement of cognitive ability and should be used independently from the vocabulary test. We ran all analyses for both the binary and the ordinal educational classification, using only the Raven test as a measure of cognitive ability. The results were very similar. Figure 10 depicts the conditional survival difference for the model without latent ability, a model with one measurement and a model with two measurements.

Even though the initial sample in 1952 was found to be representative for the Dutch population at that time, more than half of the sample is lost between 1952 and our observation period that starts in 1995. This could lead to an attrition bias, if attrition is non-random. Unfortunately, we do not have access to the original data files such that we cannot investigate attrition directly. However, Hartog (1989) investigated the non-response for the 1983 survey and found no attrition bias in a wage analysis.²² Since the sample in 1983 has been shown to be representative, we reran all analyses on just the respondents that were observed in 1983 and found no substantial changes in the results. This suggests that selective attrition does not affect our results.

The data contains information about children from different years of birth. Most of them born in earlier years had to repeat a class and their average cognitive skills are lower. This could be a potential source of selection, if staying back was due to low cognitive ability, or, worse, to health reasons. We ran a robustness check in which we excluded individuals born in 1937, 1938 and 1941. The results show that the conditional survival difference hardly deviates from the base model, if anything the difference even becomes larger. Finally, we included an indicator for the year of measurement of the education level (1983 or 1993), and again did not find any deviation. All these results are summarized in Figure 11 and the detailed estimation results are available upon request.

Finally, we varied the observed characteristics in the model. First by including additional variables among the exogenous variables such as family size, number of children, additional school characteristics (e.g. whether restricted to girls, restricted to boys, or mixed), and whether both parents were still alive. These variables were not statistically significant in any of the models, and did not alter the results. Second, we also checked robustness to excluding individuals with item non-response on some of the observed characteristics, in which case too the results remain similar.

5 Discussion

This paper estimates to what extent survival differences across educational groups are due to a ‘selection effect’ based on cognitive ability and other background variables. We extend the

²²Following Hartog, (1989) we investigated whether the attrition between 1993 and 1995 was related to observed characteristics. Literally all explanatory variables including education, family background, and intelligence were not related to attrition. The only exception was self-reported health; a worse health status increased the probability of attrition between 1993 and 1995.

structural equation model of Conti et al. (2010) to allow for a duration dependent variable and an ordinal educational choice, and estimate the model on the basis of a Dutch cohort born around 1940 for which we observe mortality between ages 55 and 75. Most important conclusion is that the selection effect based on cognitive ability is responsible for around half of the raw differences in survival. Yet, even conditional on cognitive ability and a wide range of individual characteristics, survival differences between individuals having finished only primary school and those who entered at least secondary education are still substantial, and correspond to a 4 year difference in life expectancy.

Even though we analyze mortality between ages 55 and 75 rather than self-reported health at age 30, our findings are in line with the results presented by Conti et al. (2010). Due to this striking similarity in findings, irrespective of the health measures and samples used, two tentative conclusions regarding the education-health gradient are emerging. First, at least half of the raw association between education and health is due to confounding ‘third factors’, of which cognitive ability proved very important in our analysis, while Conti et al. (2010) and Savelyev (2012) stress the importance of non-cognitive factors, in particular conscientiousness. Second, even after controlling for cognitive ability, family socioeconomic status, and a range of other background variables, education seems to remain important in determining mortality. This suggests that at least part of the educational differences in health outcomes is due to a genuine, causal effect of education on health.

However, a limitation of our data is the absence of direct measurements of non-cognitive ability. Hence, we cannot rule out that specific non-cognitive factors influence both education and health, such that our ‘conditional survival difference’ across educational groups cannot be interpreted as – and is likely to be an upper bound to – the causal effect of education on mortality. Moreover, we may overestimate the influence of cognitive ability if correlated non-cognitive abilities are omitted from the model.²³

As a starting point for sorting out this important issue, we can use the teacher’s advice regarding secondary education of the child. It is presumably a function of both the cognitive and non-cognitive abilities of the pupil. Hence, one may argue that while controlling for cognitive ability, the teacher’s advice is a proxy for non-cognitive abilities. When allowing the teacher’s advice to influence mortality directly, on top of being a determinant of educational choice, the conditional survival differences become smaller, as illustrated in Figure 11. This evidence corroborates our main conclusion that the selection effect explains *at least* half of the association between education and mortality. It also suggests that the remaining causal effect of education on mortality is likely to be smaller when taking non-cognitive abilities into account.

Another limitation is the relatively small sample size, which naturally compromised the precision of our results. Care should be taken in interpreting the exact fraction of survival differences into a selection effect and a part attributable to education, given the large uncertainty around these estimates. While the conditional survival differences across educational groups from age 73 onwards as well as the extrapolated differences in life

²³Although the literature suggests that most non-cognitive abilities are uncorrelated with IQ (Borghans et al. 2011; Savelyev, 2012).

expectancy across educational groups are statistically significant, we cannot rule out the absence of conditional differences in survival across educational groups at younger ages.

A fruitful avenue for future research would be to investigate the effect of both education and cognitive abilities on health outcomes using a more elaborate set of non-cognitive abilities. In such research the literature could benefit from our structural equation model that allows for a duration dependent variable like mortality, and an ordinal independent variable such as educational attainment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Abbring, Jaap; vandenBerg, Gerard. The Nonparametric Identification of Treatment Effects in Duration Models. *Econometrica*. 2003; 71(5):1491–1517.
- Albouy, Valerie; Lequien, Laurent. Does compulsory education lower mortality? *Journal of Health Economics*. 2008; 28(1):155–168. [PubMed: 18976824]
- Arendt, Jacob Nielsen. Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification. *Economics of Education Review*. 2005; 24(2):149–60.
- Auld, M Christopher; Sidhu, Nirmal. Schooling, Cognitive Ability and Health. *Health Economics*. 2005; 14(10):1019–1034. [PubMed: 16167318]
- Bago d’Uva, Teresa; O’Donnell, Owen; van Doorslaer, Eddy. Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*. 2008; 37(6):1375–1383. [PubMed: 18676985]
- Batty, G David; Deary, Ian J.; Gottfredson, Linda S. Premorbid (early life) IQ and Later Mortality Risk: Systematic Review. *Annals of Epidemiology*. 2007; 17(4):278–288. [PubMed: 17174570]
- Batty, G David; Wennerstad, Karin M.; Smith, George Davey; Gunnell, David; Deary, Ian; Tynelius, Per; Rasmussen, Finn. IQ in Early Adulthood and Mortality By Middle Age: Cohort Study of 1 Million Swedish Men. *Epidemiology*. 2009; 20(1):100–109. [PubMed: 19234402]
- Borghans, Lex; Golsteyn, Bart HH.; Heckman, James J.; Humphries, John E. Unpublished manuscript. University of Maastricht; 2011. IQ, Achievement, and Personality.
- Braakmann, Nils. The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England. *Journal of Health Economics*. 2011; 30(4):753–763. [PubMed: 21715033]
- Brinch, Christian N.; Galloway, Taryn A. Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences*. 2012; 109(2):425–430.
- Butler JS, Moffitt Robert. A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica*. 1982; 50(3):761–764.
- Carneiro, Pedro; Crawford, Claire; Goodman, Alissa. The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes. CEE DP. 2007:92.
- Carpenter, Patricia A.; Just, Marcel A.; Shell, Peter. What one intelligence test measures: a theoretical account of processing in the Raven Progressive Matrices Test. *Psychological Review*. 1990; 97(3): 404–431. [PubMed: 2381998]
- Anne, Case; Fertig, Angela; Paxson, Christine. The lasting impact of childhood health and circumstance. *Journal of Health Economics*. 2005; 24(2):365–389. [PubMed: 15721050]
- Centraal Bureau voor de Statistiek, CBS. Gezondheid en zorg in cijfers 2008. CBS; 2008. Hoogopgeleiden leven lang en gezond.
- Clark, Damon; Royer, Heather. The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*. 2013 forthcoming.

- Cockx, Bart; Picchio, Matteo. Are short-lived jobs stepping stones to long-lasting jobs? *Oxford Bulletin of Economics and Statistics*. 2012; 74(5):646-675.
- Conti, Gabriella; Heckman, James J. Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions. *Perspectives on Psychological Science*. 2010; 5(5):585–605. [PubMed: 21738556]
- Conti, Gabriella; Heckman, James J.; Urzua, Sergio. The Education-Health Gradient. *American Economic Review Papers and Proceedings*. 2010; 100:234–238.
- Conti, Gabriella; Heckman, James J.; Urzua, Sergio. Unpublished manuscript. University of Chicago, Department of Economics; 2011. Early Endowments, Education, and Health.
- Cramer, Jan S. Tinbergen Institute Discussion Paper 2012-070/4. 2012. Childhood Intelligence and Adult Mortality, and the Role of Socio-Economic Status.
- Cutler, David; Lleras-Muney, Adriana. Education and Health: Evaluating Theories and Evidence. In: House James, S.; Schoeni Robert, F.; Kaplan George, A.; Harold, Pollack, editors. *Making Americans Healthier: Social and Economic Policy as Health Policy*. New York: Russell Sage Foundation; 2008.
- Deary, Ian. Why do intelligent people live longer? *Nature*. 2008; 456:175–176. [PubMed: 19005537]
- Deary, Ian; Johnson, Wendy. Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*. 2010; 39:1362–1369. [PubMed: 20504860]
- Dronkers, Jaap. Bestaat er een samenhang tussen echtscheiding en intelligentie? *Mens & Maatschappij*. 2002; 77(1):25–42.
- Elbers, Chris; Lanjouw, Jean O.; Lanjouw, Peter. Micro-level estimation of poverty and inequality. *Econometrica*. 2003; 71(1):355–364.
- Elbers, Chris; Ridder, Geert. True and spurious duration dependence: The identifiability of proportional hazards models. *Review of Economic Studies*. 1982; 49:403–410.
- Fuchs, Viktor R. Time Preference and Health: An Exploratory Study. In: Fuchs, V., editor. *Economic Aspects of Health*. Chicago: The University of Chicago Press; 1982.
- Gavrilov, Leonid A.; Gavrilova, Natalia S. *The Biology of Life Span: A Quantitative Approach*. New York: Harwood Academic Publisher; 1991.
- Gottfredson, Linda. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*. 1997; 24:13–23.
- Gottfredson, Linda. Intelligence: is the Epidemiologists' elusive "Fundamental Cause" of Social Class Inequalities in Health? *Journal of Personality and Social Psychology*. 2004; 86(1):174–199. [PubMed: 14717635]
- Hartog, Joop. Survey non-response in relation to ability and family background: structure and effects on estimated earnings functions. *Applied Economics*. 1989; 21:387–395.
- Hartog, Joop; Pfann, Gerard. Vervolgonderzoek Noord-Brabantse zesdeklassers 1983, Verantwoording van hernieuwde gegevensverzameling onder Noordbrabantse zesdeklassers van 1952. Amsterdam: University of Amsterdam; 1985.
- Hartog, Joop; Oosterbeek, Hessel. Health, wealth and happiness: why pursue a higher education? *Economics of Education Review*. 1998; 17(3):245–256.
- Hartog, Joop; Jonker, Nicole; Pfann, Gerard. Documentatie Brabant data. Amsterdam: Netherlands Institute for Scientific Information Services; 2002.
- Heckman, James J.; Humphries, John E.; Veramendi, Greg; Urzua, Sergio. NBER Working Paper No. 19971. 2014. Education, Health, and Wages.
- Heckman, James J.; Singer, Burton. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*. 1984; 52:271–320.
- Kaestner, Robert; Callison, Kevin. Adolescent Cognitive and Non-Cognitive correlates of Health. *Journal of Human Capital*. 2011; 5(1):29–69.
- Lillard, Lee A. Simultaneous equations for hazards: Marriage duration and fertility timing. *Journal of Econometrics*. 1993; 56:189–217. [PubMed: 12318328]
- Lleras-Muney, Adriana. The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies*. 2005; 72:189–221.

- Mathijssen, Mathias AJM.; Sonnemans, GJM. Schoolkeuze en schoolsucces bij VHMO en ULO in Noord-Brabant. Tilburg: Zwijssen; 1958.
- Mazumder, Bhaskar. Does Education Improve Health: A Reexamination of the Evidence from Compulsory Schooling Laws. *Economic Perspectives*. 2008; 33(2)
- Mazumder, Bhaskar. The Effects of Education on Health and Mortality. *Nordic Economic Policy Review*. 2012; 2012:261–301.
- Meghir, Costas; Palme, Martin; Simeonova, Emilia. NBER Working Paper 19002. 2013. Education, Cognition and Health: Evidence from a Social Experiment.
- Murasko, Jason E. A lifecourse study on education and health: The relationship between childhood psychosocial resources and outcomes in adolescence and young adulthood. *Social Science Research*. 2007; 36(4):1348–1370.
- Oreopoulos, Philip. Estimating Average and Local Average Treatment Effects of Education when Compulsory School Laws Really Matter. *American Economic Review*. 2006; 96(1):152–175.
- Press, William H.; Flannery, Brian P.; Teukolsky, Saul A.; Vetterling, William T. *Numerical Recipes in C: The Art of Scientific Computing*. 2. Cambridge: Cambridge UP; 1993.
- Raven, John C. *Mill Hill Vocabulary Scale*. 2. London: H. K. Lewis; 1958.
- Ross, Catherine E.; Wu, Chia-ling. The links between education and health. *American Sociological Review*. 1995; 60(5):719–745.
- Savelyev, Peter A. Unpublished manuscript. Vanderbilt University, Department of Economics; 2012. Conscientiousness, Education, and Longevity of High-Ability Individuals.
- Spearman, Charles. *The Abilities of Man: Their Nature and Measurement*. New York: Macmillan; 1927.
- Van Kippersluis, Hans; O'Donnell, Owen; van Doorslaer, Eddy. Long Run Returns to Education: Does Schooling lead to an Extended Old Age? *Journal of Human Resources*. 2011; 46(4):695–721.
- Van Praag, Mirjam. Zomaar een dataset: 'Noordbrabantse zesde klassers' Een presentatie van 15 jaar onderzoek. Amsterdam: University of Amsterdam; 1992.
- Van Schellen, Marieke; Nieuwbeerta, Paul. De invloed van de militaire dienstplicht op de ontwikkeling van crimineel gedrag. *Mens & Maatschappij*. 2007; 82(1):5–27.

Highlights

- “Education and Health: The Role of Cognitive Ability” Govert Bijwaard, Hans van Kippersluis, and Justus Veenman
- We estimate the relative effects of education and cognitive ability on mortality
- We propose a structural equation model that allows for a duration outcome
- At least half of the survival gap across educational groups is due to selection
- Cognitive ability is driving most of the selection into education and mortality
- Conditional on cognitive ability, the effect of education is still substantial

Figure 1.
Kaplan-Meier Survival function by education level in two categories (top) and three categories (bottom)

Figure 2.
Relationship between cognitive ability and the binary measure for education.

Figure 3. Conditional survival differences (see equation (15)) by age and binary education category variable.

Figure 4.

Decomposition of unconditional difference in the Kaplan-Meier Survival function into conditional differences and a selection effect based on observed characteristics and cognitive ability, with binary education variable and two measurements for cognitive ability

Figure 5.

Decomposition of observed difference in the Kaplan-Meier Survival function into conditional differences and a selection effect due to observed characteristics and cognitive ability, and other selection effects based on observed characteristics only, with binary education variable and two measurements for cognitive ability (with 90% confidence intervals, below)

Figure 6. Conditional survival differences by age, ordinal education variable, two measurements for cognitive ability

Figure 7.

Decomposition of observed difference in the Kaplan-Meier Survival function into conditional differences and a selection effect based on observed characteristics and cognitive ability, with ordinal education variable and two measurements for cognitive ability (with 90% confidence intervals: lower left primary to vocational education; lower right vocational to higher education)

Figure 8.

Decomposition of observed difference in the self-reported health into conditional differences and a selection effect based on observed characteristics and cognitive ability, with binary education variable and two measurements for cognitive ability (with 90% confidence bars)

Figure 9.

Decomposition of observed difference in the Kaplan-Meier Survival function into conditional differences and a selection effect based on observed characteristics and cognitive ability, for males (left) and females (right).

Figure 10.

Conditional survival differences by age, binary education variable. Base: two measurements for cognitive ability; Model without latent skills and model with 1 measurement for cognitive ability (Raven test)

Figure 11.

Conditional survival differences by age, binary education variable using alternative models.

Base: two measurements for cognitive ability; 1983: only respondents observed in 1983;

non-cog: non-cognitive skills included; 1939-40: only respondents born in 1939-40;

edudum: dummy for measurement of education included

Table 1

Descriptive Statistics of the Brabant Data sample

Variable	Average	Standard Deviation	Number of Observations
Dependent Variables			
Mortality	0.16	0.35	2,579
Independent Variables			
Education			
Lower Education	0.14	0.34	2,537
Lower Vocational Education	0.34	0.48	2,537
At least General Secondary School	0.51	0.35	2,537
Raven p.m. test	102.04	13.28	2,579
Vocabulary test	101.42	12.87	2,579
Control Variables			
Male	0.58	0.49	2,579
Birth Rank	2.50	2.55	2,412
Family Socioeconomic Status			
Lowest	0.53	0.50	2,409
Middle	0.44	0.50	2,409
Highest	0.03	0.16	2,409
Child Works	0.28	0.45	2,256
School Religion			
Roman-Catholic	0.76	0.43	2,518
Protestant	0.19	0.40	2,518
Special	0.03	0.17	2,518
Public	0.02	0.13	2,518
Number of Teachers	6.92	2.47	2,452
Repeat			
No Repetition of Grade	0.64	0.48	2,462
Repeated Once	0.27	0.45	2,462
Repeated Twice or More	0.09	0.28	2,462
Teacher's Advice			
Continue Primary School	0.24	0.43	2,429
Lower Vocational Education	0.38	0.48	2,429
Lower Secondary Education	0.24	0.43	2,429
Higher Secondary Education	0.14	0.20	2,429
Preference of the Parents			
Work in Family Company	0.13	0.33	2,200
Paid Work without Vocational Education	0.20	0.28	2,200
Paid Work with Vocational Education	0.27	0.44	2,200
General Secondary Education	0.41	0.49	2,200

Notes: Author's calculations on the basis of the Brabant Data linked to the municipality register and the mortality register.

Table 2

Duration model - Binary education variable, two measurements for ability

Outcome	Education D	Raven Test M_1	Vocabulary Test M_2	Hazard $\lambda^{(0)}$	Hazard $\lambda^{(1)}$
Cognitive Ability					
α	0.36*** (0.09)	9.65*** (1.49)	10.37*** (1.58)	-0.33 (0.27)	-0.20 (0.13)
Constant term					
c	2.13*** (0.21)	3.59*** (0.74)	4.69*** (0.69)	-11.68*** (1.50)	-10.71*** (0.80)
a				0.11*** (0.02)	0.09*** (0.01)
Control variables					
Male	-0.25*** (0.08)	-0.93* (0.53)	-0.87* (0.49)	0.33 (0.25)	0.66*** (0.12)
Child is working - base is "No"					
Yes	-0.29*** (0.09)	-3.84*** (0.63)	-7.14*** (0.58)	0.34 (0.26)	0.15 (0.14)
Missing	-0.32*** (0.14)	-1.07 (0.90)	-2.52*** (0.83)	-0.87 (0.57)	0.12 (0.19)
Family Socioeconomic Status - base is "Low"					
Middle	0.42*** (0.10)	2.55*** (0.54)	2.24*** (0.50)	-0.34 (0.30)	0.00 (0.12)
High	0.42 (0.46)	4.16*** (1.64)	4.68*** (1.52)	-0.34 (0.30)	0.44 (0.29)
Missing	-0.54*** (0.18)	-4.39*** (1.30)	-7.63*** (1.20)	-0.67 (0.59)	0.17 (0.30)
Birthrank - base is "First"					
Second	-0.15 (0.12)	0.53 (0.79)	-0.02 (0.73)	-0.17 (0.38)	-0.00 (0.16)
Third or Fourth	-0.09 (0.11)	-0.22 (0.73)	-2.70*** (0.68)	0.09 (0.35)	-0.19 (0.16)
Fifth or higher	-0.09 (0.11)	-3.02*** (0.73)	-4.52*** (0.68)	0.09 (0.35)	-0.26* (0.16)
Missing	0.11 (0.31)	-0.63 (1.47)	0.47 (1.36)	1.13* (0.62)	-0.65* (0.34)
School religion - base is "Catholic"					
Protestant	0.31*** (0.11)	0.62 (0.68)	2.59*** (0.63)		
Other	0.42** (0.20)	5.19*** (1.13)	7.32*** (1.04)		
Number of teachers - base is "5-8 teachers"					
4	-0.16 (0.10)	-3.81*** (0.73)	-3.16*** (0.67)		

Outcome	Education <i>D</i>	Raven Test <i>M</i> ₁	Vocabulary Test <i>M</i> ₂	Hazard $\chi^{(0)}$	Hazard $\chi^{(1)}$
9 – 12	0.05 (0.10)	0.37 (0.63)	0.42 (0.59)		
Missing	0.33 (0.22)	0.81 (1.30)	0.66 (1.21)		
Teacher's advice - base is "Lower vocational school"					
Continued primary school	-0.22** (0.09)				
Lower general secondary school	0.42** (0.17)				
Higher general secondary school	0.39 (0.26)				
Missing	-0.56** (0.25)				
Repeat grade - base is "None"					
Once	-0.30*** (0.09)				
Twice or more	-0.74*** (0.12)				
Missing	0.74* (0.42)				
Preference of the parents - base is "Only vocational education"					
Work in own company	-0.78*** (0.19)				
Work without education	-1.29*** (0.19)				
Work with education	-0.88*** (0.20)				
General secondary school	-0.28 (0.18)				
Missing	-0.85*** (0.18)				

*
p-value < 0.1,

**
p-value < 0.05,

p-value < 0.01

Notes: Author's calculations on the basis of the Brabant Data linked to the municipality register and the mortality register.