Netspar
Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

The Causal Effect of Education on the Occurrence of Chronic Diseases Later in Life

Gianmaria Niccodemi

MSc 06/2017-024

University of Groningen
Faculty of Economics and Business
Master's Thesis Economics

The Causal Effect of Education on the Occurrence of Chronic Diseases Later in Life

Gianmaria Niccodemi
S3044688

Supervisors:
Dr. Viola Angelini
Dr. Govert Bijwaard

June 2017

Abstract

Objectives

Several studies have found a positive association between education and health. It is commonly assumed that a large part of this association derives from the causal effect of education on health outcomes. Confounders that affect both education choices and health, such as parental background and intelligence, may play an important role in shaping this association. Ignoring this endogeneity will bias the estimated impact of education. One solution to this problem is to explicitly modelling the interdependence of educational attainment, health outcomes, and confounders through a structural model. We use such a model to investigate the differences by education in the occurrence of chronic diseases later in life. We also determine how much of these differences is explained by education, and how much by latent intelligence and by observed characteristics.

Data

The data we use include all men who were born in The Netherlands in 1944-1947 and examined at age 18 for national conscription; a sample of 45,037 individuals is linked to medication use over the period 2006-2014. Different medications are used to identify the occurrence of chronic diseases later in life.

Methods

A three-part structural model, consisting of *(i)* a sequential probit model for the educational attainment *(ii)* a measurement system using IQ-tests to identify latent intelligen-

ce *(iii)* a probit model for medication use. Both the educational attainment model and the medication use model also depend on this latent intelligence. Based on the estimation results we derive the causal impact of education on the occurrence of chronic diseases.

Results

We find little evidence of significant causal effect of education. After controlling for observed and unobserved confounders, education significantly affects the occurrence of few chronic diseases, at some levels of education and intelligence only. A simpler non-structural model which does not control for latent intelligence leads to overestimating the causal effect of education on the occurrence of the diseases.

Contents

Chapter 1

Introduction

Several studies observe a positive association between education and health. According to Lochner [12], highly educated people may have higher income, savings and retirement benefits, and consequently high-quality health insurance and healthcare over the entire lifetime; also, they may have a balanced diet, or be more likely to avoid bad habits such as smoking or overconsumption of alcohol, eventually being more aware of the risks of unhealthy behaviours. However, whether this association derives from causation is under debate. For instance, intelligent individuals may obtain higher education levels and income, and consequently high-quality healthcare; similarly, being born in a wealthy family may have positive effects on both education and quality of the diet. In other words, confounders such as intelligence and parental background affect both educational attainment and health outcomes, playing an important role in shaping this strong association. Some papers do not take account of this endogeneity. For instance, Montez et al. [15] estimate risks of death from different chronic diseases, controlling for age, gender, and education, but not for the confounders that affect both educational attainment and risks of death. Similarly, Smith [18] estimates probit regressions on the occurrence of diabetes, controlling for education level and observed covariates such as age, weight, and smoking, but not for intelligence. Not accounting for the endogeneity eventually leads to biased estimates of the causal effect of educa-

tion. However, some strategies are possible in order to disentangle the causal effect of education on health outcomes.

One compelling option is exploiting reforms that increase the school age within a country. Examples include Clark and Royer [4], who study the effect of two reforms, in 1947 and 1974, which increased from 14 to 15 and from 15 to 16 respectively the minimum school leaving age in the United Kingdom, finding no significant causal effect of the additional education on both mortality and health outcomes. The advantage of this approach is that it is not necessary to account for any endogeneity, being the change in the schooling age exogenous.[1] However, the key drawback of the method is that the results are limited to the causal effect of one additional compulsory year within the same education level only.[2] No conclusions can be formulated on the causal effect of neither improving the education level, for example from high school to university, nor increasing the number of years of education within a different education level.

A second promising approach is modelling the interdependence of educational attainment, health outcomes, and confounders, such as latent intelligence and parental background, through a structural model. Many studies develop structural models which focus on either health outcomes early in life (see for example Conti et al. [5] and Heckman et al. [9]) or mortality and gains in life expectancy associated with higher education (see Bijwaard et al. [2, 3]). Commonly, these structural models include a measurement system which allows to identify the latent variables (intelligence and, eventually, non-cognitive skills) and consequently disentangle the causal effect of education, controlling for the same unobserved and other observed confounders.

This thesis follows the latter approach, exploiting a three-part structural model to estimate the causal effect of education on the occurrence of chronic diseases later in life. The occurrence of chronic diseases is identified from medication use in old age. This structural model allows to estimate the education choice among four different educa-

[1] In other words, one year more of education is compulsory by law. Thus, confounders such as intelligence and parental background do not play any role in increasing the education among the individuals who are affected by the reform.

[2] High school, in Clark and Royer [4].

tion levels, estimate its causal impact on the occurrence of common chronic diseases later in life, and identify and take account of the latent intelligence (and other observed factors) that affect both the education choice and the occurrence of the diseases. In particular, this three-part structural model follows a similar one developed by Bijwaard et al. [2, 3] and is composed of a sequential probit model for the educational attainment, a measurement system using IQ-tests to identify latent intelligence, and a probit model for medication use. Both the educational attainment model and the medication use model also depend on the latent intelligence. Based on the estimation results, the causal impact of education on the occurrence of chronic diseases is derived. On the one hand, the average treatment effect of obtaining the consecutive education level is computed, both for the treated (who actually obtained the consecutive education level) and the untreated individuals (who did not), and for each chronic disease. On the other hand, the differences between education levels in the probabilities of an average individual (namely, computed at the mean of the observed confounders) being affected by a certain chronic disease later in life are predicted. This is done over a significant range of the latent intelligence, allowing to understand for which levels of the latent intelligence the average individual would be significantly less or more likely to develop the disease improving the education level. Moreover, the posterior distribution of the latent intelligence conditional on the education level is computed. This result allows to estimate the percentage of a population of average individuals, with certain education, being significantly less or more likely to develop the disease by obtaining the consecutive education level. Finally, the different results are compared.

The data represent all men born in The Netherlands in 1944-1947 who were examined at 18 years old for national conscription, including variables on the education level according to the Dutch schooling system (primary, lower vocational, lower secondary, or higher), health and socio-economic status at birth or at 18 years old, and the scores of three IQ-tests which allow to identify the latent intelligence. A sample of 45,037 of these men is linked to medication use over the period 2006-2014. The chronic diseases

or conditions identified from medication use include hyperlipidemia, diabetes, COPD[3] (and bronchitis, asthma), depression (and anxiety), heartburn, hypertension, cardiac diseases, and ischemic heart disease.

The main results can be summarised as follows. After controlling for intelligence and observed characteristics, improving the education from a certain level to the consecutive one still negatively affects the occurrence of COPD, diabetes, hyperlipidemia, heartburn, and ischemic heart disease, both for the treated and the untreated individuals in the dataset. Remarkably, depression increases with higher education. The differences in the predicted probabilities of an average individual are coherent with the latter results, being significant for the same diseases and the same pair of consecutive education levels. However, they are only significant for a given interval of the latent intelligence. Generally, by obtaining the consecutive education level, the average individual would be significantly less likely to develop the disease in old age in the following cases: if intelligence is moderate to high and education is low (primary); if intelligence is around the average and education is moderately low (lower vocational); if intelligence is low to moderate and education is moderately high (lower secondary). Finally, the average treatment effects from a simple (non-structural) medication use probit model, which does not control for latent intelligence, are compared with the average treatment effects from the structural model. The comparison shows that not accounting for the confounders leads to overestimating the causal effect of education on the occurrence of the diseases. In view of all these results, the causal effect of education is hardly ever significant. Indeed, after controlling for the endogeneity the significance of the causal effect depends not only on the considered disease and education level, but also on the level of intelligence.

This thesis contributes to the literature in mainly three ways. First, little research has been conducted, through structural models, on the interdependence among education, observed and unobserved confounders, and health outcomes in old age. Thus, this re-

[3]Chronic obstructive pulmonary disease.

search aims to fill the gap in the literature, exploring the causal effect of education on the occurrence of chronic diseases or conditions later in life. Second, this research gives evidence that not accounting for the observed and unobserved confounders leads to significantly overestimating the causal effect of education. This result is valuable for future research, reinforcing the importance of using structural models, rather than simpler ones, to control for endogeneity that comes from observed and unobserved factors. Third, very little or no research on the education gradient in health at different values of latent intelligence exists to date. This thesis provides an understanding of the specific role that intelligence plays in determining the differences by education in the probability of occurrence of chronic diseases. As already mentioned, the two contributions on the role of latent intelligence are *(i)* understanding at which level of intelligence an individual may be significantly better or worse off in old age improving the education level *(ii)* estimating the percentage of the population with certain education that might be significantly better or worse off improving the education level. These are original contributions to the literature, but limited to an average (imaginary) individual and to a population of average individuals respectively. However, these results are coherent with more general results (the average treatment effects), suggesting that intelligence may play an important role in determining the significance of the educational gain not only for the average individual, but also for the individuals in the dataset.

The results suffer from some limitations. Firstly, many papers show that healthier adolescents are more likely to obtain higher education levels (see for example Jackson [11]). Unfortunately, the health status variables included in the dataset are collected at the national conscription, when the individuals are 18 years old. Therefore, the health gradient in education cannot be controlled for, since including these variables in the education choice model would lead to bias estimates of the respective parameters, due to reverse causality. Indeed, healthier individuals may perform better at school, and higher education may positively affect health at young age. A solution might be including health status variables if these are observed before the end of compulsory

school, thereby avoiding reverse causality. The second limitation is that only men are included in the dataset, restricting the results to men only. Indeed, there might be significant gender differences in the causal effect of education on health outcomes (see for instance Conti et al. [5]). Finally, mortality and occurrence of chronic diseases are not independent; therefore, there might be selective attrition due to mortality, since individuals who were more likely to develop chronic diseases were reasonably more likely to die before 2006. In other words, these individuals were more likely to be excluded from the sample linked to medication use than other individuals.

This thesis is organized as follows. Chapter 2 presents the three-part structural model, the likelihood, and the formulas to compute the causal effect of education, based on the estimates from the structural model. Chapter 3 presents data, results, and discussion of the results. Appendix A presents the treatment effects computed from a simple (non-structural) probit model; these treatment effects are compared with the respective ones from the structural model. Appendix B presents the mean of the covariates used to compute the predicted probabilities, by education level, of an average individual being affected by a certain disease in old age.

Chapter 2

The structural model

The structural model follows a similar model estimated by Bijwaard et al. [2, 3], which is composed of three parts. The first part is the education choice, the second part is the measurement system, which measures the relationship of the latent intelligence and three IQ-tests, and the third part models the occurrence of chronic diseases. The latter part differs from the model estimated by Bijwaard et al., since the final outcome is the probability of the occurrence of chronic diseases later in life rather than the mortality and survival rates. The main assumption of the structural model is that the three parts of the model are interrelated through some observed covariates at birth and over the lifespan and the latent intelligence. In other words, conditional on the observed covariates, the interdependence among the three parts of the structural model comes from the latent intelligence only. The latter jointly affects the education choice, the occurrence of chronic diseases in old age, and the IQ-tests scores, justifying the use of the structural model.

2.1 Education choice

The education choice is modelled by the sequential probit developed by Heckman et al. [9]. Each individual i chooses either to stop at a certain level of education or to

continue to the consecutive level. Four possible levels of education are considered: primary, lower vocational, lower secondary, and higher. Define the education level $d = 1, \ldots 4$, where 1 is primary, 2 is lower vocational, 3 is lower secondary, and 4 is higher. Each individual with education level $d-1$ chooses to continue to the next education level if $E_d^* > 0$, where E_d^* is the latent difference in utility between choosing the education level d and staying at level $d-1$. Therefore, a binary outcome variable E_d is defined to be equal to 1 if the individual obtained the education level d and equal to 0 otherwise, given that $E_{d-1} = 1$. It is assumed that the education choice depends on some observed covariates X and the latent intelligence θ only. Equation (1) defines the sequential probit model as a latent variable model: E_d^* is a linear combination of the observed vector of covariates X and the latent intelligence θ

$$E_d^* = X^T \beta_d + \gamma_d \theta + \epsilon_d \tag{1}$$

where $E_{d-1} = 1$ and $\epsilon_d \sim \mathcal{N}(0, 1)$. Thus, equation (2) defines the conditional probability of $E_d = 1$, given X, θ, and $E_{d-1} = 1$

$$\Pr(E_d = 1 | X, \theta, E_{d-1} = 1) = \Phi\left(X^T \beta_d + \gamma_d \theta\right) \tag{2}$$

where Φ is the CDF of the standard normal distribution.

Under the assumption that the education choice depends on X and θ only, equation (3) defines the conditional probability of an individual ending at the education level d, given X and θ

$$\Pr(E_d = 1, E_{d+1} = 0 | X, \theta) = \left\{ \prod_{h=1}^{d} \Pr(E_h = 1 | X, \theta) \right\} \times \Pr(E_{d+1} = 0 | X, \theta) \tag{3}$$

where $\Pr(E_d = 1 | X, \theta) = 1$ if $d = 1$, and $\Pr(E_{d+1} = 0 | X, \theta) = 1$ if $d = 4$. Since each individual in the dataset obtained at least the primary level, only three models need to be estimated, for $d = 2, 3, 4$ given $E_{d-1} = 1$.

2.2 Measurement system

At least three indicators of the latent intelligence are needed for identifying the measurement system (see Shipley [17, pp. 164-171]). Thus, the latter is composed of three equations, linking the scores in the IQ-test $q = 1, 2, 3$ to the latent intelligence θ and the observed covariates X. The continuous score M_q^* in the test q is not available within the dataset, but only observed in 6 different classes $c = 1, \ldots, 6$. In other words, only $M_q = 1, \ldots, 6$ is observed. Therefore, three ordered probit models need to be estimated. Equation (4) defines the latent variable M_q^* as a linear combination of the vector of observed covariates X and the latent intelligence θ

$$M_q^* = X^T \xi_q + \zeta_q \theta + \tau_q \tag{4}$$

where $\tau_q \sim \mathcal{N}(0, 1)$. Define the unknown cut points $a_{q,0}, \ldots, a_{q,6}$, where $-\infty = a_{q,0} < a_{q,1} < \ldots < a_{q,6} = +\infty$, $\forall q = 1, 2, 3$. It follows that $M_q = c$ if $a_{q,c-1} < M_q^* \leq a_{q,c}$. Thus, the conditional probability of an individual scoring $M_q = c$, given X and θ, is defined by equation (5)

$$\Pr\left(M_q = c | X, \theta\right) = \Phi\left(a_{q,c} - X^T \xi_q - \zeta_q \theta\right) - \Phi\left(a_{q,c-1} - X^T \xi_q - \zeta_q \theta\right) \tag{5}$$

where $\Phi\left(a_{q,c-1} - X^T \xi_q - \zeta_q \theta\right) = 0$ if $c = 1$, $\Phi\left(a_{q,c} - X^T \xi_q - \zeta_q \theta\right) = 1$ if $c = 6$, and Φ is the CDF of the standard normal distribution. The intercept of the model is suppressed in order to identify and estimate the parameters through the maximum likelihood method.[1] Since the intelligence is unobserved, it is necessary to establish the unit of measurement by constraining ζ_1 to one.

[1] Indeed, without imposing this restriction, the difference between the intercept and the cut point would appear in the $\Pr\left(M_q = c | X, \theta\right)$, $\forall c = 1, \ldots, 6$, and consequently in the likelihood function, since if the intercept $\xi_{q,0}$ is included, then $\Pr\left(M_q = c | X, \theta\right) = \Phi\left(\left(a_{q,c} - \xi_{q,0}\right) - X^T \xi_q - \zeta_q \theta\right) - \Phi\left(\left(a_{q,c-1} - \xi_{q,0}\right) - X^T \xi_q - \zeta_q \theta\right)$. Thus, there would not be an unique set of parameters that maximizes the likelihood function.

2.3 Probability of observing chronic diseases by education level

The occurrence of chronic diseases is modelled through a probit for each medication and education level. The medication use is observed over a span of 9 years, from 2006 to 2014, when all the individuals in the dataset are between 59 and 70 years old. Therefore, the probit allows the estimation of the probability that an individual uses a specific type of medication at some age between 59 and 70 years old. Under the assumption that there is a one-to-one correspondence between using the medication and being affected by the related disease, the probit consequently allows estimating the probability that an individual is affected by some chronic diseases in old age.

Equation (6) defines the latent utility $U^*_{m,d}$ of an individual with education level d using medication m, as a linear combination of the vector of observed covariates C and the latent intelligence θ

$$U^*_{m,d} = C^T \eta_{m,d} + \delta_{m,d}\theta + v_{m,d} \tag{6}$$

where $v_{m,d} \sim \mathcal{N}(0,1)$.[2] It is therefore defined a binary variable $U_{m,d} = 1$ if an individual with education level d used medication m at least once over the 2006-2014 period, and $U_{m,d} = 0$ otherwise. Consequently, equation (7) defines the conditional probability of an individual with education level d, being affected by the chronic disease that requires medication m, over the 2006-2014 period, given C and θ

$$\Pr\left(U_{m,d} = 1 | C, \theta\right) = \Phi\left(C^T \eta_{m,d} + \delta_{m,d}\theta\right) \tag{7}$$

where Φ is the CDF of the standard normal distribution.

[2] A slightly different set of covariates C from X is used for the medication use probit. On the other hand, the same set of observed covariates X is used for both the education choice and the measurement system. See section 3.1 for more details.

2.4 Likelihood function

To estimate the model, maximum likelihood estimation is used. Following equations (3), (5), and (7), the likelihood contribution of individual i with education level d for medication m is given by equation (8)

$$\begin{aligned}\mathcal{L}_i\left(\Psi|\theta,X,C\right) = &\left\{\prod_{c=1}^{6}\left[\Pr\left(M_1=c|X,\theta\right)\right]^{B_{ci}}\left[\Pr\left(M_2=c|X,\theta\right)\right]^{V_{ci}}\left[\Pr\left(M_3=c|X,\theta\right)\right]^{Z_{ci}}\right\}\times\\ &\times\left[\Pr\left(U_{m,d}=1|C,\theta\right)\right]^{U_{m,d}}\left[\Pr\left(U_{m,d}=0|C,\theta\right)\right]^{(1-U_{m,d})}\times\\ &\times\left\{\prod_{h=1}^{d}\Pr\left(E_h=1|X,\theta\right)\right\}\times\Pr\left(E_{d+1}=0|X,\theta\right)\end{aligned} \tag{8}$$

where $B_{ci}=1$ if $M_1=c$ and $B_{ci}=0$ otherwise, $V_{ci}=1$ if $M_2=c$ and $V_{ci}=0$ otherwise, $Z_{ci}=1$ if $M_3=c$ and $Z_{ci}=0$ otherwise, and Ψ is the set of parameters β_d, γ_d, ξ_q, ζ_q, $a_{q,1},\ldots,a_{q,5}$, $\eta_{m,d}$, $\delta_{m,d}$ to be estimated, with $q=1,2,3$.

Since the variable θ is not observed, it is necessary to assume a distribution for the intelligence and integrate it out. It is assumed that $\theta\sim\mathcal{N}\left(0,\sigma^2\right)$, where σ is the unknown standard deviation. Therefore, the likelihood contribution of individual i with education level d for medication m, not conditioning on the latent variable θ, is given by equation (9)

$$\mathcal{L}_i\left(\Psi_2|X,C\right)=\int_{-\infty}^{+\infty}\mathcal{L}_i\left(\Psi|\theta,X,C\right)f\left(\theta;0,\sigma^2\right)d\theta \tag{9}$$

where Ψ_2 includes the same set of parameters as Ψ and the standard deviation σ of θ, and $f\left(\theta;0,\sigma^2\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\left(\frac{\theta}{\sigma\sqrt{2}}\right)^2\right]$ is the PDF of the normal distribution $\mathcal{N}\left(0,\sigma^2\right)$. The maximization of the likelihood function (9) implies the calculation of an integral which does not have an analytical solution. The approximation of the solution is computed through the Gauss-Hermite quadrature. The likelihood function (9) is maximized in two steps, since only a sample of the individuals used to estimate the measurement system, the education choice, and the distribution of the latent in-

telligence θ is linked to medication use over the 2006-2014 period. In other words, the two-steps approach takes advantage of the larger dataset to compute better estimates of the measurement system, the education choice, and the distribution of the latent intelligence. The second step estimates of the medication use probit model are obtained by maximizing the likelihood (9), using the maximum likelihood estimates of the first step.

2.5 The causal effect of education

2.5.1 Conditional predicted probabilities by education level

Based on the second step estimates of the medication use probit model, the conditional probabilities by education level and intelligence can be predicted, fixing the observed covariates C included in equation (7) at certain values. In particular, the probabilities are predicted at the average of the sample linked to medication use. According to the estimated distribution of θ, the predicted probability of medication use m for education level d, at the average of the observed covariates $C = \bar{c}$, and over a relevant range of values of the latent intelligence, is defined by equation (10)

$$\Pr\left(U_{m,d} = 1 | C = \bar{c}, \theta\right) = \Phi\left(\bar{c}^T \hat{\eta}_{m,d} + \hat{\delta}_{m,d}\theta\right) \tag{10}$$

where Φ is the CDF of the standard normal distribution, $\theta \in [\theta_{min}; \theta_{max}]$, and θ_{min} and θ_{max} are arbitrarily low and high levels of intelligence respectively. That is, the latent intelligence θ varies within a range of values that may significantly represent a large part of the population, according to $\hat{\sigma}^2$. Equation (10) is computed in order to show the predicted probabilities of an average individual with different education levels being affected by a certain chronic disease in old age. This is done over relevant levels of the latent intelligence and controlling for the observed confounders. In other words, the predicted probabilities in (10) may give a first impression of the causal effect of

education on the occurrence of chronic diseases, at least for an average individual. Obviously, the formula is valid for each possible value of the observed covariates and allows to compute the predicted probability by education and intelligence, conditioning on whichever value of C.

2.5.2 Testing the differences in predicted probabilities between different education levels

According to Long [13], a $\mathcal{Z}$-test can be run and the respective confidence intervals can be computed, in order to test if the predicted probabilities by education are significantly different. Define $\Lambda_{d,h}^{(m)} = \Pr(U_{m,d} = 1|C = \bar{c}, \theta) - \Pr(U_{m,h} = 1|C = \bar{c}, \theta)$, where $d, h = 1, \ldots, 4$, $d < h$, and $\theta \in [\theta_{min}; \theta_{max}]$. The null hypothesis is $\mathrm{H}_0 : \Lambda_{d,h}^{(m)} = 0$ and the test statistic is $\mathcal{Z} = \frac{\Lambda_{d,h}^{(m)}}{\sqrt{\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right)}}$, which under the null hypothesis is asymptotically distributed as a normal. Therefore, equation (11) defines the 95% confidence intervals of the difference in the predicted probabilities of medication use m, between different education levels d and h

$$\left[\Lambda_{d,h}^{(m)} - 1.96 \times \sqrt{\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right)}; \Lambda_{d,h}^{(m)} + 1.96 \times \sqrt{\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right)}\right] \tag{11}$$

while equation (12) defines the 90% confidence intervals of the difference in the predicted probabilities of medication use m, between different education levels d and h

$$\left[\Lambda_{d,h}^{(m)} - 1.645 \times \sqrt{\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right)}; \Lambda_{d,h}^{(m)} + 1.645 \times \sqrt{\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right)}\right] \tag{12}$$

where $\mathrm{Var}\left(\Lambda_{d,h}^{(m)}\right) = \mathrm{Var}\left(\Pr(U_{m,d} = 1|C = \bar{c}, \theta)\right) + \mathrm{Var}\left(\Pr(U_{m,h} = 1|C = \bar{c}, \theta)\right)$, since the covariance term is null between two different education levels d and h.

The main purpose is to verify whether $\Lambda_{d,h}^{(m)}$ is significantly different from 0, where $d, h = 1, \ldots, 4$ and $d < h$. The confidence intervals as defined by equations (11) and

(12) vary according to the level of the latent intelligence $\theta \in [\theta_{min}; \theta_{max}]$.

The delta method allows to compute the variance of the predicted probability from the variance-covariance matrix of the estimated parameters $\hat{\beta}_{m,d}$. Generally, the variance of the conditional predicted probability given the vector of covariates B is defined by equation (13)

$$\text{Var}\left(\Pr\left(U_{m,d} = 1|B\right)\right) = \left[\phi\left(B^T\hat{\beta}_{m,d}\right)\right]^2 B^T \text{Var}\left(\hat{\beta}_{m,d}\right) B \tag{13}$$

where $\text{Var}\left(\hat{\beta}_{m,d}\right)$ is the variance-covariance matrix of the estimated parameters, B includes both the observed covariates C and the latent intelligence $\theta \in [\theta_{min}; \theta_{max}]$, and ϕ is the PDF of the standard normal distribution.

2.5.3 Posterior distribution of intelligence

Define an imaginary population of average individuals[3] split in four groups, in accordance with the respective level of education $d = 1, \ldots, 4$, where 1 is primary, 2 is lower vocational, 3 is lower secondary, and 4 is higher. Thus, it is possible to compute the percentage of each of the first three groups that, with confidence level 95%, would be less or more likely to develop the disease later in life by going from education level d to $d + 1$. In order to do so, equation (14) defines the posterior distribution of the latent intelligence $\theta \in [\theta_{min}; \theta_{max}]$, given the level of education

$$f\left(\theta|E_d = 1, E_{d+1} = 0, C = \bar{c}\right) = \frac{\Pr\left(E_d = 1, E_{d+1} = 0|\theta, C = \bar{c}\right) f\left(\theta; 0, \hat{\sigma}^2\right)}{\Pr\left(E_d = 1, E_{d+1} = 0|C = \bar{c}\right)} \tag{14}$$

where $\Pr\left(E_d = 1, E_{d+1} = 0|\theta, C = \bar{c}\right)$ is computed from equation (3) by setting the covariates that are in common between X and C at the mean value $\bar{c}$ and integrating over the distribution of the other covariates. Similarly, $\Pr\left(E_d = 1, E_{d+1} = 0|C = \bar{c}\right)$ is computed from equation (3) by setting the covariates that are in common between X and C

[3] Namely, with observed covariates $C = \bar{c}$.

at $\bar{c}$ and integrating over the distribution of the other covariates and θ.[4] Consequently, it is possible to compute the expected value of the latent intelligence within each level of education $\overline{\theta}_d$, as defined in equation (15)

$$\overline{\theta}_d = \mu_d\left(\theta | E_d = 1, E_{d+1} = 0, C = \bar{c}\right) = \int_{-\infty}^{+\infty} \theta f\left(\theta | E_d = 1, E_{d+1} = 0, C = \bar{c}\right) d\theta \tag{15}$$

where $d = 1, \ldots, 4$.

Finally, equation (16) defines the percentage of individuals of group $d = 1, 2, 3$ that would be significantly less or eventually more likely to develop a certain disease later in life by going from education level d to $d + 1$

$$\Gamma_{m,d} = \left\{ \int_{\theta_{1,m,d}}^{\theta_{2,m,d}} f\left(\theta | E_d = 1, E_{d+1} = 0, C = \bar{c}\right) d\theta \right\} \times 100 \tag{16}$$

where $\theta_{1,m,d}$ and $\theta_{2,m,d}$ are the lowest and highest values of the latent intelligence $\theta \in [\theta_{min}; \theta_{max}]$ for medication use m between which $\Lambda_{d,d+1}^{(m)} = \Pr\left(U_{m,d} = 1 | C = \bar{c}, \theta\right) - \Pr\left(U_{m,d+1} = 1 | C = \bar{c}, \theta\right)$ is significantly different from 0, at the 95% confidence level.

2.5.4 ATE, TT, and TU

Heckman et al. [10] provide estimators for the average treatment effects (ATE), the effect of treatment on the treated (TT), and the effect of treatment on the untreated (TU). The ATE is defined as the average gain in the probability of developing a disease later in life from choosing to continue to education level d, conditioning on $D \in \{d-1, d\}$, where $d = 2, 3, 4$ and D is the level of education. First of all, the predicted probabilities $\Pr\left(U_{m,d} = 1 | C = c_i, D \in \{d-1, d\}\right)$ and $\Pr\left(U_{m,d-1} = 1 | C = c_i, D \in \{d-1, d\}\right)$ are computed for each individual i, by integrating out the latent intelligence. That is, re-

[4] This is only an approximation of the posterior distribution of intelligence for a population of average individuals, according to the sample of individuals linked to medication use, since X and C are different.

gardless individual i obtained the education level d or $d-1$, the predicted probabilities forcing the observation i to have both the education levels d and $d-1$ are computed, for each level of education $d = 2, 3, 4$. Next, the same predicted probabilities are integrated over the distribution of the observed covariates, as defined by equations (17) and (18):[5]

$$\Pr\left(U_{m,d} = 1 | D \in \{d-1, d\}\right) = \frac{1}{N} \sum_{i=1}^{N} \Pr\left(U_{m,d} = 1 | C = c_i, D \in \{d-1, d\}\right) \quad (17)$$

$$\Pr\left(U_{m,d-1} = 1 | D \in \{d-1, d\}\right) = \frac{1}{N} \sum_{i=1}^{N} \Pr\left(U_{m,d-1} = 1 | C = c_i, D \in \{d-1, d\}\right) \quad (18)$$

where N is the number of observations with education level either $d-1$ or d. Finally, equation (19) defines the ATE from choosing the education level d rather than $d-1$

$$ATE_d^{(m)} = \Pr\left(U_{m,d} = 1 | D \in \{d-1, d\}\right) - \Pr\left(U_{m,d-1} = 1 | D \in \{d-1, d\}\right) \quad (19)$$

The TT and TU are very similar to the ATE. Indeed, the only difference is that rather than conditioning on $D \in \{d-1, d\}$, the treatment effects are computed only for those who actually obtained the level of education d and $d-1$ respectively, by conditioning on $D = d$ (TT) and $D = d-1$ (TU).

[5] These equations are simply the approximation (the sample analog) of the integral over the distributions of C.

Chapter 3

The structural model: estimates

3.1 Data

As mentioned in section 2.4, two different datasets are used, taking advantage of the larger one to compute better estimates of the measurement system, the education choice, and the distribution of the latent intelligence. In particular, the larger dataset includes all 408,015 men who were born in The Netherlands in 1944-1947 and examined at age 18 for compulsory national conscription; therefore, only those who emigrated before or were not alive at 18 years old are not included. The second one is a subsample of 45,037 individuals who are partially linked to medication use over the period 2006-2014, when all the individuals were between 59 and 70 years old. The subsample has been originally assembled in order to study the effect of the Dutch famine of 1944-1945 on mortality (see Ekamper et al. [7]), implying that the Western famine-exposed region is oversampled. Indeed, it includes all 25,283 men of the larger dataset who were born between November 1944 and March 1946 in the most famine-exposed cities in The Netherlands, during the *Hongerwinter*.[1] The rest of the data is composed by a random sample of 10,667 individuals who were born in the same cities but before November 1944 or after March 1946, and a random sample of 9,087 individuals who

[1] According to Ekamper et al. [7], these cities are Amsterdam, Haarlem, Rotterdam, The Hague, Leiden, and Utrecht.

were born in a different part of The Netherlands in 1944-1947. From both the dataset and its subsample, the individuals who did not take all the three IQ-tests or attended special schools for disabled or illiterate are excluded. From the subsample of 45,037 individuals only, individuals who died or emigrated before 2006 are excluded, as well as those for whom medication use is unknown. After these exclusions, the final dataset amounts to 368,809 and its subsample to 33,428 individuals. It should be noted that mortality and occurrence of chronic diseases are not independent, since people who were more likely to develop chronic diseases were reasonably more likely to die before 2006 (and thus to be excluded from the sample linked to medication use) than other individuals. Therefore, the estimates of the causal effect of education may be affected by selective attrition due to mortality.[2] However, this bias should not be high since the individuals are firstly observed in 2006, at the relatively young ages of 59-62 years old.

Similar to Doornbos and Kromhout [6], four possible levels of education $d = 1, 2, 3, 4$ are defined, where $d = 1$ is primary education (six years of schooling, from 6 to 12 years old), $d = 2$ is lower vocational education (eight years of schooling), $d = 3$ is lower secondary education (ten years of schooling), and $d = 4$ is higher education, which includes intermediate vocational education, general secondary education, and higher non-university or university education (at least twelve years of schooling). Each individual in the dataset obtained at least the education level $d = 1$, therefore three models need to be estimated, for $d = 2, 3, 4$, forming the education choice model (see equation (3)). The dependent variables of the three models are three binary outcome variables E_d equal to 1 if the individual obtained the education level d and equal to 0 otherwise, where $d = 2, 3, 4$ and $E_{d-1} = 1$.

The dependent variables of the three ordered probit models, forming the measurement system (see paragraph 2.2), are the scores in the three IQ-tests $q = 1, 2, 3$, where $q = 1$ is the Raven progressive matrices test, $q = 2$ is an arithmetic test, and $q = 3$ is a language test. These tests represent the measures of the latent intelligence θ. Each test score is

[2] In other words, the estimates may differ if the individuals that died before 2006 were still alive.

observed in six ordered classes $c = 1, \ldots, 6$, where 1 and 6 are the lowest and highest class respectively. Table 1 shows the distribution of the three tests scores by education level, among the 368,809 individuals in the original dataset. It can be seen that the IQ-tests scores increase with education, suggesting that the distribution of the latent intelligence may differ conditioning on the education level.

Table 1: Distribution of IQ tests scores by education, 368,809 ind.

	Primary	Lower vocational	Lower secondary	Higher
Raven:				
1 (lowest)	12.1%	2.7%	1.2%	0.3%
2	26.3 %	10.0%	3.7 %	0.9%
3	24.1%	17.7%	9.3%	3.5%
4	21.2%	26.8%	20.4%	11.4%
5	13.8%	30.6%	38.5%	37.9%
6 (highest)	2.6 %	12.2%	27.0%	46.1%
Arith:				
1 (lowest)	19.8%	2.7%	0.3%	0.1%
2	44.3%	16.2 %	2.9%	0.4%
3	24.4%	23.8 %	9.8%	1.4%
4	9.4%	28.2%	27.8%	8.0%
5	1.9%	23.1%	40.7%	38.1%
6 (highest)	0.2%	6.1%	18.6%	52.1%
Lang:				
1 (lowest)	13.6%	1.7%	0.2%	0.0%
2	38.2%	14.8%	1.2%	0.2%
3	28.3%	28.6%	5.4%	0.7%
4	16.0%	33.5%	20.7%	6.6%
5	3.8%	18.3%	49.6%	43.4%
6 (highest)	0.2%	3.2 %	22.9%	49.1%

The final outcome is the medication use probit model by education level. The dependent variable is a binary variable $U_{m,d} = 1$ if an individual with education level d used the medication m at least once over the 2006-2014 period, and $U_{m,d} = 0$ otherwise. Eight chronic diseases or conditions are identified from medication use: hyperlipidemia, diabetes, COPD,[3] depression, heartburn, hypertension, cardiac diseases,

[3]Chronic obstructive pulmonary disease (COPD) is an obstructive lung disease, mainly caused by tobacco smoking.

and ischemic heart disease, where the latter is a particular case of cardiac diseases.[4] The main assumption is that there is a one-to-one correspondence between using the prescribed medication m (the variable observed in the dataset) and being affected by the related disease. The medication use is observed on annual basis and in ATC-code[5] with three levels of classification: the first indicates the anatomical main group, the second the therapeutic subgroup, and the third the pharmacological subgroup.

Table 2: Chronic diseases and relative medication use

Diseases or conditions	Medication use (ATC-code)
Hyperlipidemia	C10
Diabetes	A10
COPD (and bronchitis, asthma)	R03
Depression (and anxiety)	N05B, N06A
Heartburn	A02A, A02B
Hypertension	C02, C03A, C07, C08, C09A,B
Cardiac diseases	C01, C03C
Ischemic heart disease	C01D

Table 3: Medication use by education, 33,428 ind., 2006-2014

	Primary	Lower vocational	Lower secondary	Higher
Hyperlipidemia	54.6%	49.4%	45.3%	36.9%
Diabetes	22.6%	18.3%	16.5%	11.1%
COPD (and bronchitis, asthma)	30.7%	24.1%	21.2%	18.6%
Depression (and anxiety)	30.0%	24.5%	23.6%	22.0%
Heartburn	61.3%	56.3%	50.5%	45.2%
Hypertension	60.1%	57.0%	54.8%	48.5%
Cardiac diseases	29.5%	24.4%	22.3%	17.9%
Ischemic heart disease	18.3%	14.5%	12.8%	9.0%
Number of individuals	4,665	12,447	10,989	5,327

According to Halfon et al. [8] and van Ooijen et al. [16], each of the relevant chronic diseases and conditions are uniquely linked to one or more medications. These diseases and conditions and the relative medications in ATC-code are listed in Table 2; the percentage of the 33,428 individuals that used each medication at least once, over

[4]The ischemic heart disease (IHD) is also known as coronary artery disease (CAD) and is the main group of diseases within the general group of cardiac diseases.

[5]Anatomical Therapeutic Chemical Classification System.

Table 4: Medication use in subsample (S) and The Netherlands (NL), in 2007 and 2014

Medication use (ATC-code)	S: 2007, age 60-63	NL: 2007, age 60-65	S: 2014, age 67-70	NL: 2014, age 65-70
C10 (Hyperlipidemia)	27.4%	27.2%	40.8%	38.0%
A10 (Diabetes)	10.0%	10.2%	15.5%	14.0%
R03 (COPD)	8.9%	9.3%	11.9%	11.7%
N05B (Depression)	8.7%	8.0%	2.0%	1.9%
N06A	5.3%	5.3%	5.5%	5.6%
A02A (Heartburn)	0.0%	0.0%	0.0%	0.0%
A02B	18.8%	19.0%	25.2%	23.6%
C02 (Hypertension)	0.8%	0.8%	1.1%	1.0%
C03A	5.6%	5.6%	10.0%	9.1%
C07	20.5%	20.5%	27.8%	25.3%
C08	9.7%	9.6%	15.9%	14.5%
C09A	13.3%	13.1%	20.0%	18.2%
C09B	2.6%	2.5%	3.6%	3.0%
C01 (Cardiac diseases)	6.3%	6.2%	8.4%	7.2%
C03C	2.7%	2.9%	4.6%	3.9%
C01D (Ischemic heart disease)	4.3%	4.4%	5.3%	4.7%

Source for The Netherlands: CBS

the period 2006-2014 and by education level, is shown in Table 3. The latter descriptive statistics may appear unnaturally high and, consequently, the data linked to medication use may seem not representative of the Dutch population. However, it should be considered that the period of observation of each individual is 9 years, conditional on surviving.[6] To demonstrate that the data well represent the whole Dutch population, Table 4 shows the descriptive statistics of medication use in 2007 and in 2014 only, when all the individuals are 60-63 and 67-70 years old respectively, and the descriptive statistics of the whole Netherlands for slightly different age groups, namely 60-65 and 65-70 years old respectively (source for The Netherlands: CBS website). It can be seen that the percentages of the Dutch population and its sample using each medication are very similar, both in 2007 and in 2014. The little differences may depend on the different definition of the age groups.

As already mentioned in section 2.4, the likelihood (9) is maximized in two steps: the first step includes the measurement system, the education choice, and the distribution of the latent intelligence, for which all the 368,809 observations are used. The second

[6]That is, the percentage of people who use a certain medication at least once over 9 years is obviously higher than the percentage of the same people using the same medication at least once over one year only.

Table 5: Characteristics X at age 18 by education, 368,809 ind.

	Primary	Lower vocational	Lower secondary	Higher
Average height (cm)	175.8	177.1	178.1	179.1
Average birth order	3.0	2.7	2.4	2.2
Average family size	5.0	4.6	4.0	3.9
Father's occupation:				
Professional	7.2%	8.6%	17.3%	37.7%
White collar	16.3%	23.4%	37.5%	38.8%
Farm owner	7.0%	14.4%	6.6%	5.1%
Skilled	37.0%	31.1%	22.5%	9.7%
Unskilled	26.6%	17.7%	11.8%	4.7%
Unknown	5.9%	4.8%	4.3%	4.1%
Urban status of place of birth:				
Selected city	33.6%	28.8%	38.0%	42.5%
Urban	21.3%	19.4%	21.3%	21.9%
Urbanized rural	21.2%	21.6%	18.6%	17.6%
Rural	21.0%	27.3%	18.2%	13.7%
Unknown	2.9%	2.9%	3.8%	4.4%
Religion:				
Catholic	47.3%	41.9%	38.0%	39.0%
Protestant	34.6%	42.3%	43.2%	41.2%
Other	0.5%	0.4%	0.7%	0.9%
None	17.6%	15.4%	18.0%	19.0%
Famine exposure:				
PN	9.6%	9.9%	10.1%	10.1%
T3	10.3%	10.5%	10.6%	9.9%
T2	9.9%	9.6%	9.2%	8.3%
T1	8.7%	8.2%	7.7%	7.6%
PC	9.8%	10.1%	10.1%	10.0%
No Exposure	66.9%	66.7%	66.9%	68.0%
Region of birth:				
West	43.3%	41.3%	48.1%	49.6%
South	26.8%	23.3%	20.3%	21.6%
East	17.4%	19.6%	16.0%	14.4%
North	12.4%	15.7%	15.6%	14.4%
Year of birth:				
1944	23.3%	23.8%	24.5%	23.7%
1945	22.9%	22.6%	22.2%	21.2%
1946	29.9%	30.2%	30.5%	30.8%
1947	23.9%	23.3%	22.8%	24.4%
Number of individuals	56,052	154,240	107,859	50,658

Table 6: Characteristics C at age 18 by education, 33,428 ind. linked to medication use

	Primary	Lower vocational	Lower secondary	Higher
Average height (cm)	176.0	177.4	178.1	179.3
Average BMI	21.8	21.7	21.5	21.4
Father's occupation:				
Professional	8.9%	10.0%	16.8%	38.0%
White collar	19.5%	29.4%	43.3%	43.9%
Farm owner	3.0%	5.7%	2.2%	1.7%
Skilled	38.5%	33.8%	23.4%	9.3%
Unskilled	22.7%	15.2%	9.5%	3.4%
Unknown	7.4%	5.8%	4.8%	3.6%
Urban status of place of birth:				
Selected city	81.1%	78.7%	86.4%	87.8%
Urban	6.4%	6.2%	4.7%	4.9%
Non-urban	12.5%	15.2%	8.9%	7.3%
Famine exposure:				
PN	16.7%	17.5%	19.7%	19.6%
T3	16.7%	17.9%	19.3%	18.4%
T2	15.9%	14.8%	15.6%	15.0%
T1	10.5%	9.2%	10.8%	11.7%
PC	16.1%	15.6%	18.8%	19.4%
No Exposure	45.8%	46.8%	39.9%	40.1%
General health fitness:				
Fit	83.9%	85.9%	83.6%	83.3%
Lower than fit	16.1%	14.1%	16.4%	16.7%
General psychological fitness:				
Fit	65.1%	80.1%	81.1%	81.9%
Lower than fit	34.9%	19.9%	18.9%	18.1%
Year of birth:				
1944	16.1%	18.4%	17.1%	17.6%
1945	40.1%	39.3%	41.2%	40.3%
1946	29.6%	29.5%	29.9%	29.7%
1947	14.3%	12.7%	11.8%	12.4%
Number of individuals	4,665	12,447	10,989	5,327

step includes the medication use model, using the estimates of the first step and the 33,428 individuals linked to medication use. The first and second steps are computed through two different vectors of observed covariates, which are X and C respectively. X includes height, birth order, family size, father's occupation, urban status of place of

birth, region of birth, famine exposure, self-reported religion, year of birth, and month of birth. C includes father's occupation, urban status of place of birth, height, body mass index (BMI),[7] famine exposure, year of birth, and the binary variables general health fitness and psychological fitness.[8] Unfortunately, BMI, general health fitness, and general psychological fitness are observed at 18 years old and thus cannot be included in the vector X, since there might be reverse causality between education choice and health status.[9] On the contrary, height and famine exposure can be included in both datasets, because the former is mostly determined by genetic factors and the latter is a prenatal condition. In particular, height is reasonably likely to negatively affect the occurrence of diseases later in life. On the other hand, Ekamper et al. [7] show that, *ceteris paribus*, being famine exposed during the first trimester of pregnancy and in the postnatal period increases by 12% and 8% the mortality rate respectively. Thus, it is important to control for the famine exposure, since this may affect the education choice, the IQ-tests scores, and the occurrence of diseases. According to Ekamper et al. [7], the variable famine exposure is composed by categories which overlap. PN (postnatal exposure) includes births between November 1, 1944 and March 31, 1945; T3 (third trimester of pregnancy) includes births between February 1, 1945 and June 30, 1945; T2 (second trimester of pregnancy) includes births between May 1, 1945 and September 30, 1945; T1 (first trimester of pregnancy) includes births between August 1, 1945 and December 31, 1945; PC (post-conception) includes births between November 1, 1945 and March 31, 1946. Thus, there is no unique reference category: PN, T3, T2, T1, and PC are binary variables equal to one if the individual was famine exposed in the respective period of pregnancy, and equal to 0 otherwise.

The other variables in common between X and C are the most likely to have a significant effect on health outcomes in old age, by affecting the whole lifespan: indeed, father's occupation and urban status of place of birth may affect social conditions,

[7] BMI is defined as the mass divided by the square of the height.

[8] General health fitness and psychological fitness are equal to 1 if the individual was evaluated as "fit general health" and "fit psychological health" respectively, at the examination for the national conscription, and 0 if less than fit.

[9] Obviously, these variables can be included in vector X if measured before the end of primary school.

wealth, and type of occupation of the individuals, and consequently health outcomes, from birth to old age; in this respect, there is no reason to include variables such as religion and birth order in the vector C. Region of birth may be optionally included, however controlling for the similar variable of urban status of place of birth appears to be sufficient. In particular, Bijwaard et al. [3] define the latter in five categories, as follows: selected city (cities with more than 100,000 inhabitants), urban (cities with less than 100,000 inhabitants), urbanized rural (rural communities with less than 20% of farming population), rural (rural communities with more than 20% of farming population), and unknown. Urban status of place of birth included in vector C and in vector X are slightly different, since in the former the categories urbanized rural, rural, and unknown are grouped together, forming the category non-urban, due to lack of observations in the subsample linked to medication use. Father's occupation includes four different categories, ordered from the highest to the lowest level of remunerations and responsibilities: professional, white collar, skilled, and unskilled; the fifth category includes individuals whose father's occupation is unknown.

Table 5 shows the average height, birth order, and family size and the distribution of the other regressors included in vector X by education level,[10] for the 368,809 individuals in the original dataset. Table 6 shows the average BMI and height and the distribution of the other regressors included in vector C by education level, for the 33,428 individuals in the subsample linked to medication use. In regressions, both height and BMI are mean centered, and 1 is subtracted from birth order.[11] Finally, as already mentioned, in the subsample linked to medication use the famine-exposed cities are oversampled, implying that the observed percentages of both selected city and PN, T3, T2, T1, PC are much higher than the respective ones in the larger dataset.

The use of the three-part structural model to explain the causal effect of education is justified by looking at the descriptive statistics. As can be seen from Table 3, there is a substantial difference in medication use by education level, up to 8.4 percent-

[10] Except for month of birth.

[11] In other words, birth order equal to 0 means that the individual is the first born.

age points from lower secondary to higher education (hyperlipidemia), 5.8 percentage points from lower vocational to lower secondary education (heartburn), and 6.6 percentage points from primary to lower vocational education (COPD). However, not accounting for the confounders (the observed covariates and the latent intelligence) may lead to overestimating the effect of education on the occurrence of chronic diseases. Indeed, Table 3 does not reflect the causal effect of education only, since the scores in the three IQ-tests increase among those who have higher education levels (Table 1); moreover, those who obtained higher education levels usually feature characteristics which are reasonably likely to decrease the probability of developing chronic diseases or conditions later in life: for instance, the average height and the percentage of individuals whose father's occupation is professional increase, and the percentage of individuals whose father's occupation is unskilled decreases with higher levels of education (Tables 5 and 6). Since both the intelligence and the observed variables are likely to affect the education choice as well as the occurrence of chronic diseases or conditions, the three-part structural model is necessary to disentangle the causal effect of education.

3.2 First step estimates

Table 7 shows the estimates of the first step, using the dataset of the 368,809 individuals. The reference category of each categorical variable is defined within the table. It can be seen that the latent intelligence θ positively affects the probability of going from education level d to $d+1, \forall d = 1, 2, 3$. Also, higher intelligence leads to higher scores in the IQ-tests. The estimated variance of the latent intelligence (not reported in Table 7) is $\hat{\sigma}^2 = 0.776$. The latent variable should be seen as innate intelligence, namely intelligence after controlling for other factors that may affect IQ-tests scores as well. Indeed, it can be seen from Table 7 that some covariates such as height, father's occupation, urban status of place of birth, region of birth, religion, and so on significantly affect the

IQ-tests scores. Overall, the signs of the estimated coefficients are in line with expectations, since, for instance, both a father who is professional and being born in a selected city increase the probabilities of higher IQ-tests scores and education, with respect to the reference categories (white collar and urban respectively). On the other hand, both unskilled and rural have a negative effect on the same probabilities. Moreover, the negative effects of both birth order and family size on education are coherent with the literature (see for example Bagger et al. [1]), reflecting the presence of a quality (of education) and quantity (of children) trade off. As regards the famine exposure, T1 negatively affects the probability of going from primary to lower vocational education, but does not affect the probability of going from lower vocational to lower secondary and from lower secondary to higher, nor the IQ-tests scores. Similarly, T2 has some negative effects on the probability of going from primary to lower vocational education and it also negatively affects two out of three IQ-tests scores. PC positively affects both the probability of going from primary to lower vocational education and every IQ-tests scores, and PN positively affects the Raven scores.

3.3 Second step estimates

Using the first step estimates (the estimated variance of the intelligence $\hat{\sigma}^2 = 0.776$ and the estimated parameters of the education choice and the measurement system) the second step estimates are computed and reported in Tables 8-15, maximizing the likelihood function (9) for each medication. Overall, the estimated coefficient of the latent intelligence θ, as expected, negatively affects the probability of the occurrence of the diseases, being highly significant (at the 95% confidence level or above) for the individuals with medium levels of education ($d = 2, 3$), but occasionally less or not significant at the extreme levels of education ($d = 1, 4$). In particular, the latent intelligence θ has a strongly significant and negative effect on the occurrence of each chronic disease for individuals with lower vocational education. Similarly, θ negatively and

Table 7: First step estimates

	(Educ 12)	(Educ 23)	(Educ 34)	(Raven)	(Arith)	(Lang)
Constant	2.762***	0.755***	−0.923***			
	(109.24)	(40.08)	(−38.65)			
Intelligence	1.469***	1.265***	1.151***	1(*constr.*)	1.901***	1.858***
	(169.07)	(177.77)	(115.32)		(239.83)	(220.79)
Height	0.036***	0.029***	0.031***	0.028***	0.045***	0.045***
	(50.81)	(50.67)	(41.84)	(69.84)	(77.29)	(78.22)
Birth order	−0.054***	−0.067***	−0.070***	−0.051***	−0.084***	−0.108***
	(−18.93)	(−26.62)	(−20.67)	(−30.48)	(−34.45)	(−44.77)
Family size	−0.096***	−0.064***	−0.042***	−0.022***	−0.057***	−0.055***
	(−40.18)	(−31.54)	(−15.54)	(−15.79)	(−28.12)	(−27.72)
Reference category: White collar						
Professional	0.246***	0.485***	0.719***	0.208***	0.394***	0.384***
	(15.54)	(46.17)	(64.43)	(28.01)	(36.63)	(36.14)
Farm owner	−0.078***	−0.922***	−0.524***	−0.542***	−0.643***	−0.801***
	(−4.52)	(−69.10)	(−27.20)	(−59.73)	(−48.37)	(−60.95)
Skilled	−0.837***	−0.964***	−1.000***	−0.465***	−0.929***	−0.999***
	(−71.52)	(−105.06)	(−74.22)	(−74.18)	(−99.40)	(−108.12)
Unskilled	−1.085***	−1.067***	−1.105***	−0.604***	−1.171***	−1.162***
	(−82.38)	(−96.99)	(−63.62)	(−82.45)	(−107.04)	(−108.08)
Unknown	−0.669***	−0.652***	−0.363***	−0.343***	−0.773***	−0.752***
	(−33.77)	(−39.82)	(−17.20)	(−30.22)	(−46.82)	(−46.24)
Reference category: Urban						
Selected city	−0.014	0.103***	0.081***	0.082***	0.028**	0.071***
	(−1.18)	(10.68)	(6.88)	(12.21)	(2.89)	(7.42)
Urbanized	0.099***	−0.098***	−0.071***	−0.116***	−0.116***	−0.115***
	(7.67)	(−9.35)	(−5.25)	(−16.01)	(−11.04)	(−11.05)
Rural	0.104***	−0.254***	−0.242***	−0.186***	−0.207***	−0.194***
	(7.79)	(−23.37)	(−16.65)	(−24.90)	(−19.06)	(−18.11)
Unknown	0.184***	0.257***	0.264***	−0.071***	0.087***	−0.245***
	(6.85)	(12.38)	(10.37)	(−4.92)	(4.14)	(−11.80)
Reference category: West						
North	0.175***	0.022*	−0.016	−0.159***	0.068***	0.054***
	(12.37)	(1.96)	(−1.11)	(−20.78)	(6.09)	(4.96)
South	0.025*	−0.000	0.134***	0.057***	0.253***	0.130***
	(2.08)	(−0.00)	(10.88)	(8.40)	(25.69)	(13.41)
East	0.040**	−0.130***	−0.067***	−0.160***	−0.044***	−0.117***
	(3.22)	(−13.09)	(−5.19)	(−23.32)	(−4.46)	(−11.94)
Famine exposure						
PN	0.010	0.015	0.012	0.026*	0.015	0.017
	(0.48)	(0.93)	(0.60)	(2.28)	(0.88)	(1.06)
T3	−0.003	0.021	−0.015	0.010	0.018	0.001
	(−0.12)	(0.93)	(−0.53)	(0.64)	(0.78)	(0.05)
T2	−0.049*	−0.034	−0.028	−0.022	−0.071***	−0.077***
	(−2.25)	(−1.92)	(−1.22)	(−1.82)	(−4.01)	(−4.40)
T1	−0.090**	0.014	−0.013	0.013	−0.017	0.048
	(−2.81)	(0.53)	(−0.39)	(0.70)	(−0.64)	(1.85)
PC	0.050**	0.022	0.001	0.047***	0.051***	0.056***
	(2.67)	(1.43)	(0.04)	(4.52)	(3.39)	(3.76)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: First step estimates

	(Educ 12)	(Educ 23)	(Educ 34)	(Raven)	(Arith)	(Lang)
Reference category: January						
February	−0.033	−0.043*	0.014	−0.021	−0.046**	−0.008
	(−1.54)	(−2.46)	(0.63)	(−1.69)	(−2.60)	(−0.47)
March	−0.020	−0.067***	0.003	−0.002	−0.045**	−0.029
	(−0.93)	(−3.95)	(0.14)	(−0.15)	(−2.67)	(−1.72)
April	−0.017	−0.055**	−0.007	0.019	−0.036*	−0.040*
	(−0.75)	(−3.10)	(−0.32)	(1.54)	(−1.99)	(−2.24)
May	0.019	−0.033	0.005	0.041***	−0.008	0.006
	(0.87)	(−1.81)	(0.21)	(3.30)	(−0.44)	(0.33)
June	−0.003	−0.056**	−0.011	0.055***	−0.026	−0.000
	(−0.13)	(−3.07)	(4.35)	(−0.48)	(−1.44)	(−0.02)
July	0.005	−0.042*	−0.008	0.058***	−0.006	0.004
	(0.24)	(−2.42)	(−0.38)	(4.82)	(−0.34)	(0.26)
August	0.018	−0.077***	−0.032	0.049***	−0.023	−0.014
	(0.80)	(−4.13)	(−1.34)	(3.78)	(−1.23)	(−0.76)
September	0.016	−0.074***	0.014	0.055***	0.003	0.006
	(0.70)	(−3.97)	(0.58)	(4.27)	(0.14)	(0.31)
October	0.045	−0.031	0.110***	−0.020	−0.029	−0.129***
	(1.93)	(−1.67)	(4.66)	(−1.58)	(−1.51)	(−6.93)
November	0.002	−0.030	0.097***	−0.052***	−0.061***	−0.133***
	(0.09)	(−1.67)	(4.19)	(−4.12)	(−3.35)	(−7.39)
December	0.048*	−0.014	0.124***	−0.022	−0.025	−0.091***
	(2.15)	(−0.76)	(5.46)	(−1.74)	(−1.38)	(−5.08)
Reference category: Protestant						
Catholic	−0.081***	0.044***	0.140***	0.045***	−0.013	−0.114***
	(−7.68)	(5.21)	(12.98)	(7.57)	(−1.46)	(−13.56)
None	−0.190***	−0.079***	−0.029*	0.007	−0.153***	−0.153***
	(−15.18)	(−7.94)	(−2.34)	(1.00)	(−15.31)	(−15.57)
Other	−0.115*	0.414***	0.259***	−0.124***	−0.122**	−0.065
	(−2.09)	(9.22)	(5.14)	(−3.98)	(−2.66)	(−1.43)
Reference category: 1944						
1945	0.004	−0.062**	−0.022	0.037*	−0.051*	0.050*
	(0.14)	(−2.62)	(−0.74)	(2.24)	(−2.17)	(2.12)
1946	−0.036**	−0.100***	−0.040**	0.003	−0.179***	−0.051***
	(−2.92)	(−10.25)	(−3.19)	(0.48)	(−18.16)	(−5.25)
1947	−0.035**	−0.170***	−0.092***	−0.404***	−0.501***	−0.530***
	(−2.80)	(−16.97)	(−7.25)	(−58.27)	(−49.57)	(−53.24)
Cut point 1				−3.168***	−4.771***	−5.193***
				(−220.69)	(−210.44)	(−224.59)
Cut point 2				−2.198***	−3.022***	−3.391***
				(−161.51)	(−147.66)	(−165.28)
Cut point 3				−1.432***	−1.880***	−2.119***
				(−107.44)	(−95.86)	(−108.72)
Cut point 4				−0.577***	−0.616***	−0.781***
				(−43.81)	(−32.19)	(−41.33)
Cut point 5				0.668***	1.101***	1.044***
				(50.66)	(57.02)	(54.91)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

significantly affects at the 95% confidence level (or above) the probability of the occurrence of all the diseases except for ischemic heart disease, for individuals with lower secondary education. For individuals with higher education, θ does not significantly affect at the 95% confidence level the probability of diabetes, COPD, cardiac diseases, and ischemic heart disease. Similarly, for individuals with primary education, it does not significantly affect the probability of COPD, depression, and ischemic heart disease. Overall, these first results seem to suggest that a higher intelligence is likely to significantly contribute to the prevention of chronic diseases later in life for the individuals with medium education levels (lower vocational and lower secondary); on the other hand, the significance of θ for the individuals with primary or higher education depends on the type of chronic disease.

The famine exposure has rarely a remarkable effect on the probability. When significant, the effects of PN and T1 are negative, while the effects of T3 and T2 are positive. PC and T2 are hardly ever significant. The most relevant positive effect of famine exposure, in terms of number of significant famine-exposure categories, is estimated in the diabetes regression. Indeed, the occurrence of diabetes for individuals with lower vocational, lower secondary, and higher education levels is positively affected by T3, T2, and PC respectively. This evidence is coherent with the literature that observes an increasing probability in the risk of type II diabetes after famine during gestation (see for instance Lumey et al. [14]).

Finally, the signs of the significant parameters of the other observed covariates are generally in line with expectations. Firstly, a larger height significantly reduces and a higher BMI (at 18 years old) significantly increases the probability of each disease, for at least one education level, except for cardiac diseases and depression respectively. Similarly, lower than fit general health (at 18 years old) significantly increases the probability of each disease except for hyperlipidemia and ischemic heart disease, for at least one education level. Low psychological fitness (at 18 years old) has, overall, less remarkable effects, being significant (and positive) only for diabetes, COPD, and

Table 8: Second step estimates: probit hyperlipidemia

Hyperlipidemia	Primary	Lower vocational	Lower secondary	Higher
Constant	−0.046	0.025	−0.080	−0.066
	(−0.51)	(0.52)	(−1.62)	(−0.81)
Intelligence	−0.077*	−0.075***	−0.085***	−0.087*
	(−2.09)	(−3.67)	(−3.48)	(−2.25)
BMI	0.006	0.021***	0.027***	0.027**
	(0.73)	(3.90)	(4.85)	(3.11)
Height	−0.012***	−0.014***	−0.019***	−0.020***
	(−4.08)	(−7.71)	(−9.51)	(−6.62)
Lower than fit health	0.045	0.017	0.055	−0.082
	(0.89)	(0.53)	(1.66)	(−1.68)
Lower than fit psych	−0.042	−0.048	−0.037	−0.011
	(−1.03)	(−1.65)	(−1.16)	(−0.23)
Reference category: White collar				
Professional	−0.062	−0.093*	0.014	−0.017
	(−0.83)	(−2.25)	(0.39)	(−0.43)
Farm owner	0.108	−0.183***	−0.209*	0.218
	(0.90)	(−3.29)	(−2.32)	(1.51)
Skilled	0.128*	0.053	0.055	0.090
	(2.41)	(1.80)	(1.67)	(1.37)
Unskilled	0.112	0.056	0.073	0.144
	(1.89)	(1.51)	(1.62)	(1.44)
Unknown	−0.001	0.100	0.174**	0.196*
	(−0.02)	(1.94)	(2.97)	(2.05)
Reference category: Selected city				
Non-urban	−0.020	−0.044	−0.022	−0.057
	(−0.30)	(−1.13)	(−0.45)	(−0.74)
Urban	0.003	−0.0390	0.006	−0.104
	(0.03)	(−0.77)	(0.11)	(−1.20)
Reference category: 1945				
1944	0.084	−0.021	0.024	−0.046
	(1.17)	(−0.50)	(0.49)	(−0.65)
1946	−0.005	−0.054	−0.122*	−0.163*
	(−0.06)	(−1.20)	(−2.37)	(−2.14)
1947	−0.031	−0.174***	−0.143*	−0.300***
	(−0.39)	(−3.51)	(−2.54)	(−3.58)
Famine exposure				
PN	−0.036	−0.110**	0.005	−0.105
	(−0.55)	(−2.82)	(0.12)	(−1.75)
T3	0.039	0.092*	0.020	−0.029
	(0.56)	(2.10)	(0.43)	(−0.41)
T2	0.014	−0.056	0.065	−0.083
	(0.20)	(−1.29)	(1.44)	(−1.26)
T1	−0.078	0.023	−0.005	−0.141
	(−0.94)	(0.43)	(−0.10)	(−1.71)
PC	−0.044	−0.068	0.035	−0.038
	(−0.69)	(−1.72)	(0.87)	(−0.65)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Second step estimates: probit diabetes

Diabetes	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.646^{***}	-1.016^{***}	-1.005^{***}	-1.085^{***}
	(-6.40)	(-18.41)	(-17.08)	(-10.31)
Intelligence	-0.114^{**}	-0.130^{***}	-0.128^{***}	-0.073
	(-2.78)	(-5.44)	(-4.38)	(-1.42)
BMI	0.039^{***}	0.048^{***}	0.061^{***}	0.069^{***}
	(4.56)	(8.02)	(9.52)	(6.51)
Height	-0.000	-0.007^{**}	-0.014^{***}	-0.012^{**}
	(-0.05)	(-3.12)	(-6.06)	(-3.20)
Lower than fit health	0.046	0.052	0.120^{**}	0.125^{*}
	(0.81)	(1.37)	(3.09)	(2.04)
Lower than fit psych	-0.021	0.045	0.076^{*}	0.080
	(-0.46)	(1.33)	(2.04)	(1.33)
Reference category: White collar				
Professional	-0.108	-0.063	-0.060	0.010
	(-1.29)	(-1.28)	(-1.41)	(0.19)
Farm owner	-0.225	-0.117	-0.050	-0.020
	(-1.57)	(-1.71)	(-0.45)	(-0.10)
Skilled	-0.039	0.049	0.098^{*}	0.106
	(-0.67)	(1.43)	(2.52)	(1.25)
Unskilled	0.045	0.083	0.125^{*}	0.067
	(0.69)	(1.95)	(2.37)	(0.52)
Unknown	-0.109	0.010	0.116	0.220
	(-1.21)	(0.16)	(1.70)	(1.87)
Reference category: Selected city				
Non-urban	-0.231^{**}	-0.084	-0.015	-0.223^{*}
	(-3.07)	(-1.84)	(-0.25)	(-2.01)
Urban	-0.093	0.053	0.041	0.082
	(-1.02)	(0.91)	(0.57)	(0.75)
Reference category: 1945				
1944	-0.089	0.081	0.001	-0.166
	(-1.13)	(1.59)	(0.02)	(-1.80)
1946	-0.218^{**}	0.047	-0.073	-0.212^{*}
	(-2.66)	(0.88)	(-1.18)	(-2.13)
1947	-0.322^{***}	-0.022	-0.084	-0.195
	(-3.61)	(-0.38)	(-1.25)	(-1.78)
Famine exposure				
PN	-0.079	-0.029	0.084	0.036
	(-1.11)	(-0.65)	(1.73)	(0.46)
T3	-0.053	0.158^{**}	-0.020	-0.034
	(-0.69)	(3.11)	(-0.36)	(-0.38)
T2	-0.079	0.000	0.121^{*}	-0.097
	(-1.04)	(0.01)	(2.29)	(-1.12)
T1	-0.077	0.065	0.002	-0.147
	(-0.85)	(1.05)	(0.03)	(-1.36)
PC	-0.106	-0.123^{**}	-0.000	0.154^{*}
	(-1.48)	(-2.64)	(-0.01)	(2.02)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Second step estimates: probit COPD

COPD	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.599^{***}	-0.775^{***}	-0.826^{***}	-0.772^{***}
	(6.26)	(-15.06)	(-14.77)	(-8.42)
Intelligence	-0.063	-0.103^{***}	-0.132^{***}	-0.065
	(-1.65)	(-4.60)	(-4.81)	(-1.47)
BMI	-0.001	0.015^{**}	0.013^{*}	0.017
	(-0.12)	(2.61)	(2.13)	(1.74)
Height	-0.003	-0.004^{*}	-0.003	-0.009^{**}
	(-0.87)	(-2.10)	(-1.67)	(-2.83)
Lower than fit health	0.114^{*}	0.205^{***}	0.210^{***}	0.276^{***}
	(2.16)	(5.91)	(5.84)	(5.29)
Lower than fit psych	0.023	0.083^{**}	0.010	0.039
	(0.54)	(2.65)	(0.29)	(0.74)
Reference category: White collar				
Professional	0.079	0.070	0.027	-0.079
	(1.01)	(1.55)	(0.70)	(-1.71)
Farm owner	-0.046	0.086	-0.071	0.121
	(-0.36)	(1.42)	(-0.69)	(0.76)
Skilled	0.059	0.020	0.096^{**}	-0.043
	(1.06)	(0.63)	(2.61)	(-0.58)
Unskilled	0.100	0.114^{**}	0.139^{**}	-0.116
	(1.61)	(2.85)	(2.78)	(-0.98)
Unknown	0.029	0.150^{**}	0.116	-0.039
	(0.35)	(2.72)	(1.79)	(-0.36)
Reference category: Selected city				
Non-urban	-0.109	-0.113^{**}	-0.024	0.022
	(-1.59)	(-2.65)	(-0.45)	(0.25)
Urban	-0.098	-0.191^{***}	-0.198^{**}	-0.028
	(-1.15)	(-3.34)	(-2.76)	(-0.28)
Reference category: 1945				
1944	0.019	0.033	0.047	-0.104
	(0.25)	(0.70)	(0.86)	(-1.28)
1946	-0.029	-0.042	-0.045	-0.016
	(-0.37)	(-0.85)	(-0.78)	(-0.18)
1947	-0.075	-0.074	-0.021	-0.189^{*}
	(-0.90)	(-1.35)	(-0.33)	(-2.00)
Famine exposure				
PN	-0.053	-0.128^{**}	-0.051	0.028
	(-0.77)	(-2.98)	(-1.11)	(0.41)
T3	0.011	0.038	0.004	-0.020
	(0.15)	(0.79)	(0.07)	(-0.26)
T2	-0.078	0.078	0.024	-0.011
	(-1.06)	(1.68)	(0.48)	(-0.15)
T1	0.055	-0.067	-0.013	-0.088
	(0.63)	(-1.15)	(-0.20)	(-0.95)
PC	-0.058	0.001	0.052	-0.079
	(-0.88)	(0.02)	(1.15)	(-1.19)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Second step estimates: probit depression

Depression	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.605^{***}	-0.765^{***}	-0.636^{***}	-0.479^{***}
	(-6.31)	(-14.90)	(-11.80)	(-5.43)
Intelligence	-0.068	-0.125^{***}	-0.124^{***}	-0.169^{***}
	(-1.76)	(-5.64)	(-4.61)	(-3.95)
BMI	0.001	-0.005	-0.002	0.000
	(0.07)	(-0.85)	(-0.36)	(0.01)
Height	-0.000	-0.002	-0.005^{*}	-0.003
	(-0.00)	(-1.01)	(-2.48)	(-0.80)
Lower than fit health	0.059	0.057	0.097^{**}	0.085
	(1.11)	(1.62)	(2.71)	(1.63)
Lower than fit psych	0.044	0.134^{***}	0.144^{***}	0.087
	(1.05)	(4.35)	(4.25)	(1.73)
Reference category: White collar				
Professional	0.040	-0.018	-0.007	0.016
	(0.52)	(-0.40)	(-0.18)	(0.37)
Farm owner	-0.128	-0.032	0.061	0.045
	(-1.00)	(-0.53)	(0.63)	(0.28)
Skilled	0.039	0.048	0.033	-0.005
	(0.70)	(1.49)	(0.93)	(-0.07)
Unskilled	-0.053	0.068	0.027	0.191
	(-0.85)	(1.70)	(0.55)	(1.76)
Unknown	0.055	0.095	-0.007	-0.057
	(0.65)	(1.72)	(-0.11)	(-0.52)
Reference category: Selected city				
Non-urban	-0.025	-0.080	0.023	-0.151
	(-0.36)	(-1.88)	(0.44)	(-1.77)
Urban	0.045	-0.070	-0.058	-0.050
	(0.54)	(-1.27)	(-0.87)	(-0.53)
Reference category: 1945				
1944	0.004	-0.030	-0.126^{*}	-0.171^{*}
	(0.05)	(-0.62)	(-2.39)	(-2.19)
1946	0.029	0.026	-0.106	-0.204^{*}
	(0.37)	(0.52)	(-1.89)	(-2.45)
1947	-0.025	0.008	-0.130^{*}	-0.151
	(-0.30)	(0.15)	(-2.13)	(-1.67)
Famine exposure				
PN	-0.148^{*}	-0.029	-0.047	-0.079
	(-2.11)	(-0.67)	(-1.05)	(-1.19)
T3	-0.028	0.039	-0.031	-0.114
	(-0.37)	(0.81)	(-0.62)	(-1.51)
T2	0.059	0.005	-0.057	-0.085
	(0.80)	(0.10)	(-1.15)	(-1.17)
T1	-0.012	-0.054	-0.150^{*}	-0.275^{**}
	(-0.14)	(-0.92)	(-2.43)	(-3.02)
PC	-0.060	-0.072	0.046	0.080
	(-0.90)	(-1.66)	(1.04)	(1.25)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Second step estimates: probit heartburn

Heartburn	Primary	Lower vocational	Lower secondary	Higher
Constant	0.130	0.095*	0.038	0.183*
	(1.41)	(2.02)	(0.77)	(2.28)
Intelligence	−0.108**	−0.121***	−0.137***	−0.136***
	(−2.89)	(−5.95)	(−5.62)	(−3.57)
BMI	0.007	0.010	0.008	0.020*
	(0.91)	(1.83)	(1.38)	(2.39)
Height	−0.005	−0.004	−0.004	−0.009**
	(−1.80)	(−1.91)	(−1.95)	(−3.09)
Lower than fit health	0.144**	0.083*	0.107**	0.042
	(2.77)	(2.49)	(3.23)	(0.89)
Lower than fit psych	−0.074	0.016	0.037	−0.045
	(−1.82)	(0.54)	(1.19)	(−0.98)
Reference category: White collar				
Professional	0.075	−0.044	−0.038	−0.019
	(0.98)	(−1.07)	(−1.09)	(−0.48)
Farm owner	−0.107	−0.046	−0.038	0.112
	(−0.89)	(−0.84)	(−0.44)	(0.79)
Skilled	0.042	0.001	0.061	−0.001
	(0.79)	(0.05)	(1.88)	(−0.01)
Unskilled	0.133*	0.078*	0.064	0.114
	(2.20)	(2.10)	(1.43)	(1.16)
Unknown	0.007	0.088	0.120*	−0.093
	(0.09)	(1.69)	(2.05)	(−0.98)
Reference category: Selected city				
Non-urban	−0.051	−0.017	0.067	0.026
	(−0.77)	(−0.43)	(1.38)	(0.34)
Urban	0.008	−0.124*	−0.057	−0.144
	(0.10)	(−2.46)	(−0.96)	(−1.70)
Reference category: 1945				
1944	0.032	0.047	0.016	−0.128
	(0.45)	(1.08)	(0.32)	(−1.82)
1946	−0.015	0.022	−0.057	−0.210**
	(−0.20)	(0.48)	(−1.11)	(−2.80)
1947	−0.059	−0.023	−0.081	−0.168*
	(−0.74)	(−0.46)	(−1.45)	(-2.06)
Famine exposure				
PN	−0.078	0.012	−0.013	−0.074
	(−1.18)	(0.31)	(−0.33)	(−1.25)
T3	−0.014	−0.007	−0.017	−0.091
	(−0.20)	(−0.16)	(−0.36)	(−1.33)
T2	0.101	0.026	0.001	−0.066
	(1.42)	(0.61)	(0.02)	(−1.02)
T1	−0.038	−0.011	−0.035	−0.188*
	(−0.45)	(−0.20)	(−0.62)	(−2.32)
PC	0.013	−0.005	0.035	0.060
	(0.20)	(−0.13)	(0.88)	(1.04)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Second step estimates: probit hypertension

Hypertension	Primary	Lower vocational	Lower secondary	Higher
Constant	0.136	0.207***	0.236***	0.128
	(1.48)	(4.40)	(4.74)	(1.59)
Intelligence	−0.100**	−0.082***	−0.066**	−0.110**
	(−2.68)	(−4.03)	(−2.69)	(−2.91)
BMI	0.032***	0.042***	0.049***	0.056***
	(4.02)	(7.84)	(8.59)	(6.52)
Height	−0.001	−0.006**	−0.007***	−0.011***
	(−0.26)	(−3.19)	(−3.57)	(−3.89)
Lower than fit health	0.075	0.035	0.071*	0.040
	(1.45)	(1.07)	(2.14)	(0.85)
Lower than fit psych	−0.051	0.032	0.009	0.044
	(−1.25)	(1.09)	(0.29)	(0.95)
Reference category: White collar				
Professional	0.021	−0.085*	−0.015	−0.094*
	(0.28)	(−2.04)	(−0.44)	(−2.37)
Farm owner	−0.103	−0.138*	−0.128	0.004
	(−0.86)	(−2.50)	(−1.46)	(0.03)
Skilled	0.059	−0.003	0.070*	0.111
	(1.11)	(−0.10)	(2.14)	(1.71)
Unskilled	0.072	0.006	0.033	0.118
	(1.21)	(0.17)	(0.74)	(1.19)
Unknown	0.017	0.049	0.173**	0.107
	(0.21)	(0.95)	(2.93)	(1.13)
Reference category: Selected city				
Non-urban	0.029	−0.075	−0.051	0.038
	(0.43)	(−1.94)	(−1.05)	(0.50)
Urban	−0.062	−0.111*	0.054	−0.114
	(−0.76)	(−2.19)	(0.91)	(−1.35)
Reference category: 1945				
1944	0.012	0.033	−0.056	0.038
	(0.17)	(0.76)	(−1.16)	(0.54)
1946	−0.061	−0.090*	−0.224***	−0.053
	(−0.82)	(−1.98)	(−4.36)	(−0.70)
1947	−0.145	−0.091	−0.159**	−0.190*
	(−1.81)	(−1.83)	(−2.83)	(-2.32)
Famine exposure				
PN	−0.009	−0.038	0.003	−0.042
	(−0.13)	(−0.96)	(0.06)	(−0.71)
T3	0.131	0.098*	−0.027	0.041
	(1.82)	(2.21)	(−0.58)	(0.60)
T2	0.064	−0.017	0.001	0.019
	(0.91)	(−0.38)	(0.02)	(0.30)
T1	−0.098	−0.061	−0.124*	−0.036
	(−1.17)	(−1.14)	(−2.21)	(−0.44)
PC	−0.051	−0.044	−0.036	−0.020
	(−0.79)	(−1.11)	(−0.90)	(−0.34)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: Second step estimates: probit cardiac diseases

Cardiac diseases	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.595^{***}	-0.674^{***}	-0.683^{***}	-0.868^{***}
	(-6.19)	(-13.23)	(-12.52)	(-9.33)
Intelligence	-0.084^{*}	-0.084^{***}	-0.067^{*}	-0.081
	(-2.16)	(-3.76)	(-2.46)	(-1.82)
BMI	0.039^{***}	0.037^{***}	0.042^{***}	0.038^{***}
	(4.76)	(6.63)	(6.93)	(3.94)
Height	-0.001	-0.002	-0.003	0.002
	(-0.22)	(-0.99)	(-1.30)	(0.71)
Lower than fit health	0.029	0.064	0.104^{**}	0.031
	(0.55)	(1.80)	(2.87)	(0.57)
Lower than fit psych	0.032	0.053	0.032	0.023
	(0.74)	(1.69)	(0.91)	(0.43)
Reference category: White collar				
Professional	-0.163^{*}	0.016	-0.004	0.002
	(-2.03)	(0.36)	(-0.09)	(0.03)
Farm owner	-0.189	-0.098	-0.192	0.341^{*}
	(-1.45)	(-1.58)	(-1.83)	(2.23)
Skilled	-0.009	0.014	0.045	-0.030
	(-0.17)	(0.44)	(1.24)	(-0.39)
Unskilled	-0.038	0.125^{**}	0.083	0.094
	(-0.60)	(3.14)	(1.68)	(0.83)
Unknown	-0.019	0.121^{*}	0.164^{**}	0.088
	(-0.23)	(2.21)	(2.60)	(0.81)
Reference category: Selected city				
Non-urban	-0.022	-0.087^{*}	-0.048	0.118
	(-0.32)	(-2.05)	(-0.88)	(1.37)
Urban	0.145	-0.082	-0.000	-0.034
	(1.71)	(-1.48)	(-0.00)	(-0.33)
Reference category: 1945				
1944	0.010	-0.012	-0.047	0.060
	(0.14)	(-0.25)	(-0.89)	(0.74)
1946	-0.117	-0.062	-0.161^{**}	-0.120
	(-1.50)	(-1.27)	(-2.84)	(-1.38)
1947	-0.216^{*}	-0.155^{**}	-0.188^{**}	-0.126
	(-2.55)	(-2.84)	(-3.01)	(-1.32)
Famine exposure				
PN	-0.022	-0.136^{**}	-0.042	0.030
	(-0.32)	(-3.19)	(-0.94)	(0.44)
T3	0.106	0.022	-0.069	0.050
	(1.44)	(0.46)	(-1.36)	(0.63)
T2	0.002	-0.034	0.015	-0.030
	(0.03)	(-0.72)	(0.31)	(-0.39)
T1	-0.051	-0.018	-0.085	0.032
	(-0.58)	(-0.31)	(-1.38)	(0.34)
PC	0.096	-0.079	0.037	0.073
	(1.42)	(-1.82)	(0.81)	(1.09)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Second step estimates: probit ischemic heart disease

Ischemic heart disease	Primary	Lower vocational	Lower secondary	Higher
Constant	-1.049^{***}	-1.071^{***}	-1.118^{***}	-1.320^{***}
	(-9.84)	(-18.53)	(-17.84)	(-11.72)
Intelligence	-0.035	-0.088^{***}	-0.059	-0.088
	(-0.81)	(-3.50)	(-1.90)	(-1.64)
BMI	0.015	0.026^{***}	0.019^{**}	0.034^{**}
	(1.66)	(4.05)	(2.71)	(3.00)
Height	-0.010^{**}	-0.0143^{***}	-0.015^{***}	-0.008
	(-2.97)	(-6.22)	(-6.02)	(-1.84)
Lower than fit health	-0.081	0.003	0.035	-0.033
	(-1.35)	(0.07)	(0.84)	(-0.49)
Lower than fit psych	0.018	0.027	0.056	0.102
	(0.39)	(0.78)	(1.41)	(1.63)
Reference category: White collar				
Professional	-0.052	0.021	-0.007	0.067
	(-0.58)	(0.40)	(-0.16)	(1.21)
Farm owner	-0.050	-0.138	-0.176	0.176
	(-0.34)	(-1.87)	(-1.41)	(0.94)
Skilled	0.094	0.037	0.034	-0.037
	(1.53)	(1.03)	(0.81)	(-0.39)
Unskilled	0.043	0.136^{**}	0.124^{*}	-0.189
	(0.62)	(3.05)	(2.23)	(-1.19)
Unknown	-0.001	0.153^{*}	0.186^{**}	0.211
	(-0.01)	(2.50)	(2.65)	(1.69)
Reference category: Selected city				
Non-urban	-0.001	-0.103^{*}	-0.020	0.075
	(-0.01)	(-2.12)	(-0.31)	(0.74)
Urban	0.127	-0.042	0.033	-0.137
	(1.37)	(-0.67)	(0.43)	(-1.09)
Reference category: 1945				
1944	0.083	-0.023	-0.032	0.099
	(1.00)	(-0.43)	(-0.52)	(0.99)
1946	-0.050	-0.052	-0.121	0.008
	(-0.58)	(-0.93)	(-1.86)	(0.08)
1947	-0.083	-0.134^{*}	-0.117	0.011
	(-0.89)	(-2.15)	(-1.63)	(0.10)
Famine exposure				
PN	0.022	-0.108^{*}	-0.030	-0.061
	(0.29)	(-2.25)	(-0.57)	(-0.74)
T3	0.114	0.028	0.027	0.142
	(1.41)	(0.53)	(0.46)	(1.46)
T2	0.091	0.011	0.012	-0.043
	(1.13)	(0.21)	(0.22)	(-0.48)
T1	-0.020	-0.047	-0.045	0.111
	(-0.21)	(-0.72)	(-0.63)	(0.97)
PC	0.106	-0.002	0.093	-0.098
	(1.41)	(-0.05)	(1.80)	(-1.20)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

depression, and for individuals with lower vocational or lower secondary education. When significant, being one or two years younger (namely, being born in 1946 or 1947 with respect to 1945) decreases the probability of each disease; however, being one year older (being born in 1944) is hardly ever significant. With respect to selected cities, less urbanized places of birth (urban and non-urban) occasionally reduce the probabilities of occurrence of chronic diseases for certain education levels, but never significantly increase it. Finally, father's occupation categories, when significant, have the expected signs: with respect to white collar, more remunerative professions (professional) decrease the probability of chronic diseases, while less remunerative professions (skilled and unskilled) increase it.

3.4 The causal effect of education

3.4.1 Conditional predicted probabilities by education level

Following equation (10), the predicted probabilities are computed for each education level, at the mean of the observed covariates of the sample linked to medication use, and over a relevant range of values of the latent variable. According to the estimated variance of the latent intelligence $\hat{\sigma}^2$, the probabilities are predicted for $\theta \in [-2;2]$. Indeed, the estimated standard deviation of θ is $\hat{\sigma} = \sqrt{0.776} = 0.881$ and by setting $\theta_{min} = -2$ and $\theta_{max} = 2$, the probability of the intelligence being within the interval $[-2;2]$ is $\Pr(-2 < \theta < 2) = \Pr\left(\frac{-2}{0.881} < z < \frac{2}{0.881}\right)$, where $z \sim \mathcal{N}(0,1)$. Consequently, $\Pr(-2 < \theta < 2) = \Phi\left(\frac{2}{0.881}\right) - \Phi\left(\frac{-2}{0.881}\right) = 0.977$, where Φ is the CDF of the standard normal distribution. This means that around 98% of the 368,809 individuals in the original dataset have $\theta \in [-2;2]$, since $\theta \sim \mathcal{N}(0, 0.776)$.

These predicted probabilities by education and intelligence are represented in a graph for each medication use. In other words, Figures 1-8 represent $\Pr(U_{m,d} = 1 | C = \bar{c}, \theta)$ for each medication m, education level $d = 1,2,3,4$, and $\theta \in [-2;2]$. Obviously, it is

possible to compute the predicted probabilities at other levels of the observed covariates, in order to predict the probability of a specific individual with a certain education level developing a certain disease later in life.

From Figures 1-8, it can be seen that the predicted probability of each chronic disease, for all education levels, decreases as the latent intelligence increases.[12] In addition, conditions such as heartburn, hypertension, and hyperlipidemia are more likely to occur than illnesses such as COPD (and bronchitis, asthma), diabetes, and cardiac diseases (including ischemic heart disease), holding constant the latent intelligence and the education level. The reason is simply that the former are common conditions among the elderly, while the latter are more serious chronic diseases. Depression (and anxiety) is a special case, being more related to mental rather than physical health. The chance of medications for depression being used seems to be in-between the probability of the common conditions and the chronic illnesses.

The probabilities of the average individual with primary and higher education developing the same disease are generally the highest and the lowest respectively, holding θ constant. Exceptions include depression and hypertension. According to Figure 6, the predicted probabilities of hypertension overlap at $\theta = -0.3$ and thus do not seem to significantly differ by education, holding θ constant; on the other hand, according to Figure 4, for all $\theta \in [-2; -0.1]$, the probability of the average individual using medications for depression is the highest if the individual obtained the higher education level; moreover, for all $\theta \in [-0.1; 1.1]$, the predicted probability is still higher than the ones of the average individual with lower vocational and lower secondary education. Differently from the predictions for primary and higher education, the expected probabilities of all the diseases for the medium education levels $d = 2, 3$ (lower secondary and lower vocational) are always close together and occasionally overlap, suggesting that there is hardly ever significant difference between them.

[12] As already seen in Tables 8-15, the latent intelligence is occasionally not significant; however, the sign of the estimated parameter is always negative, implying that the predicted probability always decreases if the intelligence increases.

Figure 1: Probability of hyperlipidemia by education and intelligence

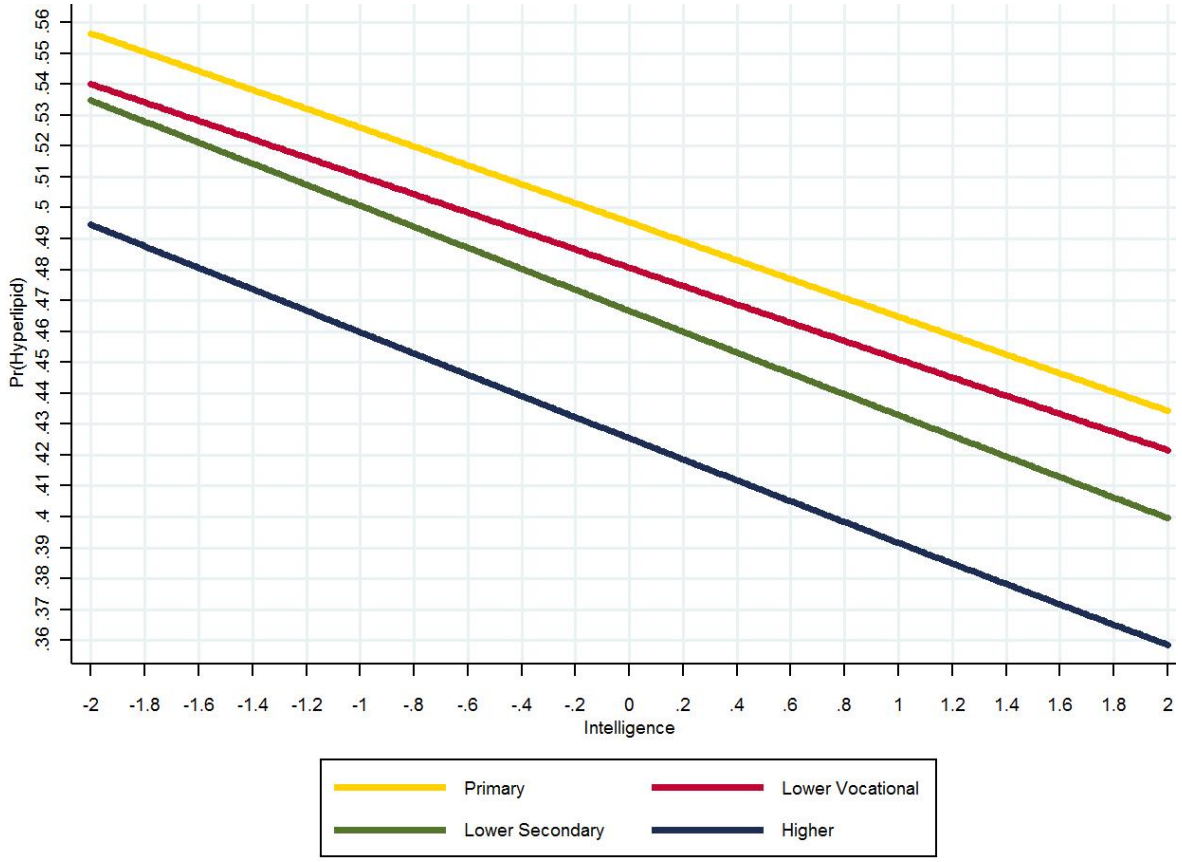


Figure 2: Probability of diabetes by education and intelligence

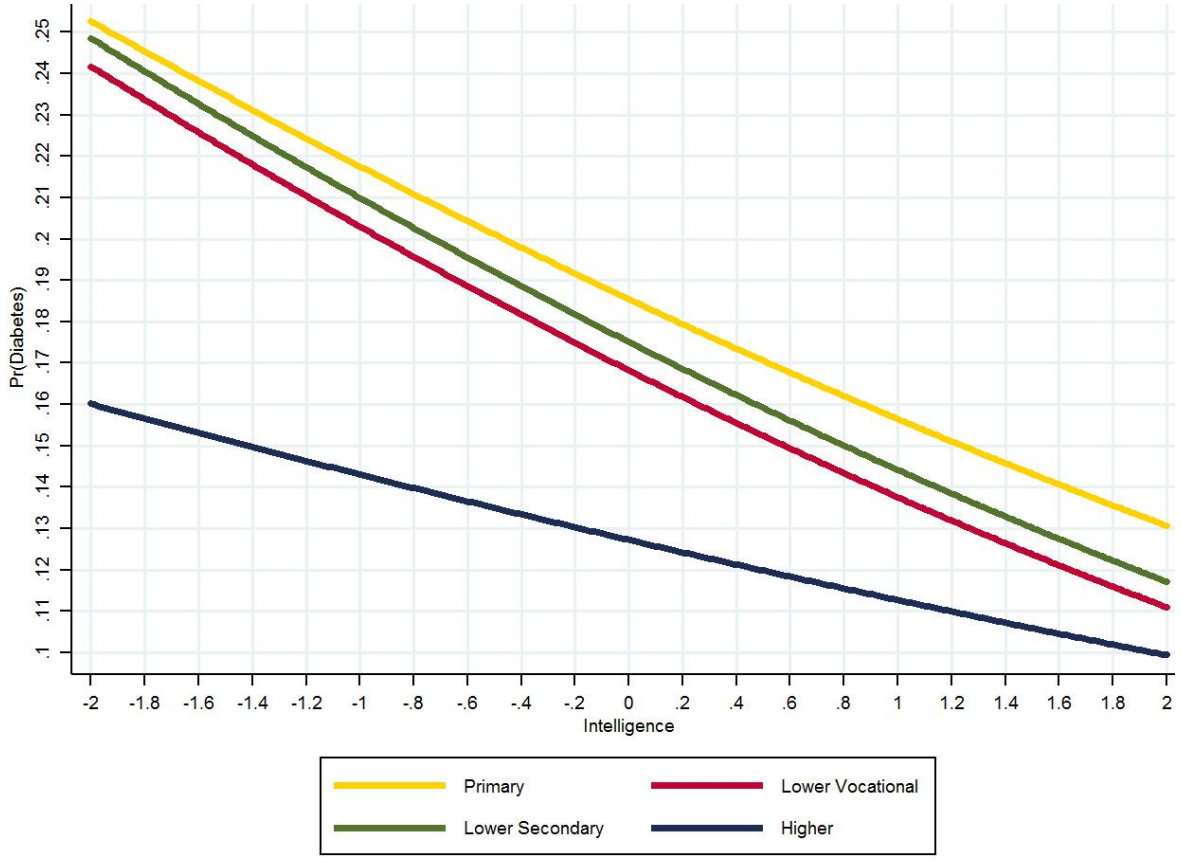


Figure 3: Probability of COPD by education and intelligence

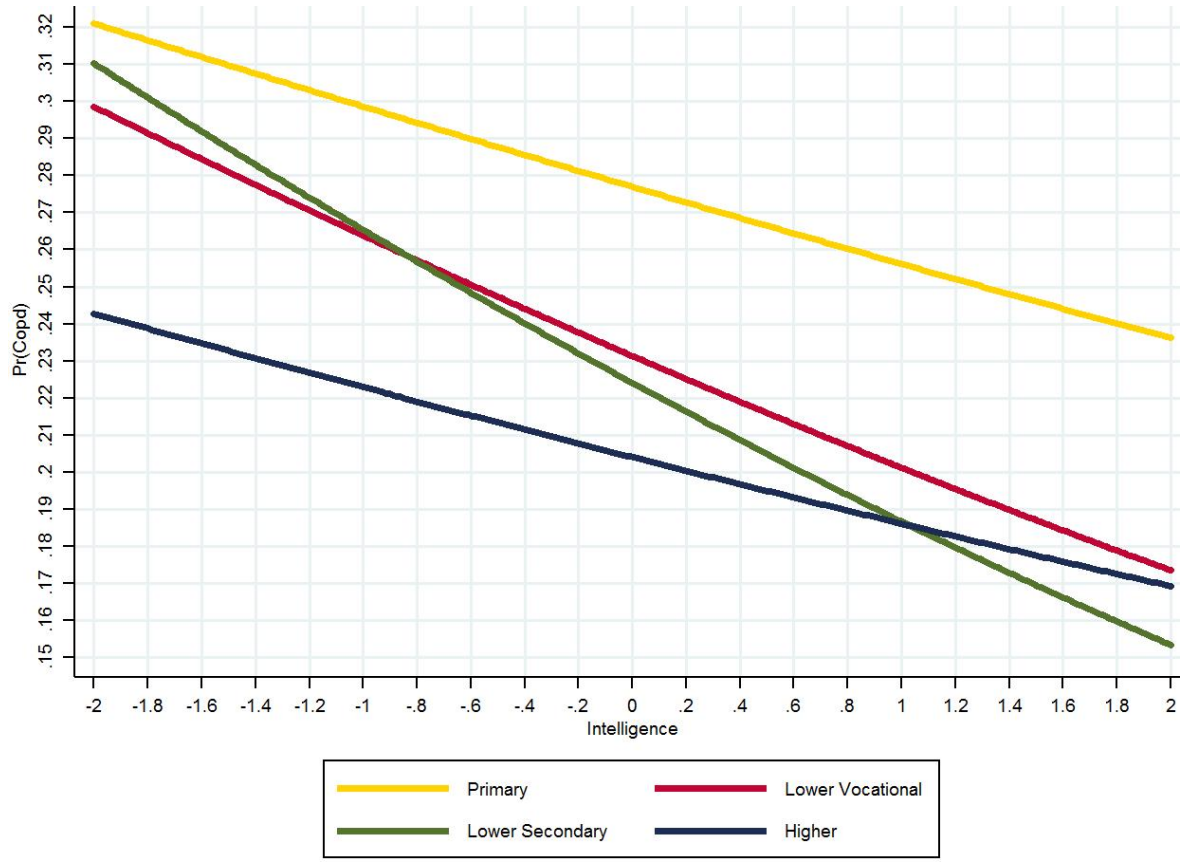


Figure 4: Probability of depression by education and intelligence

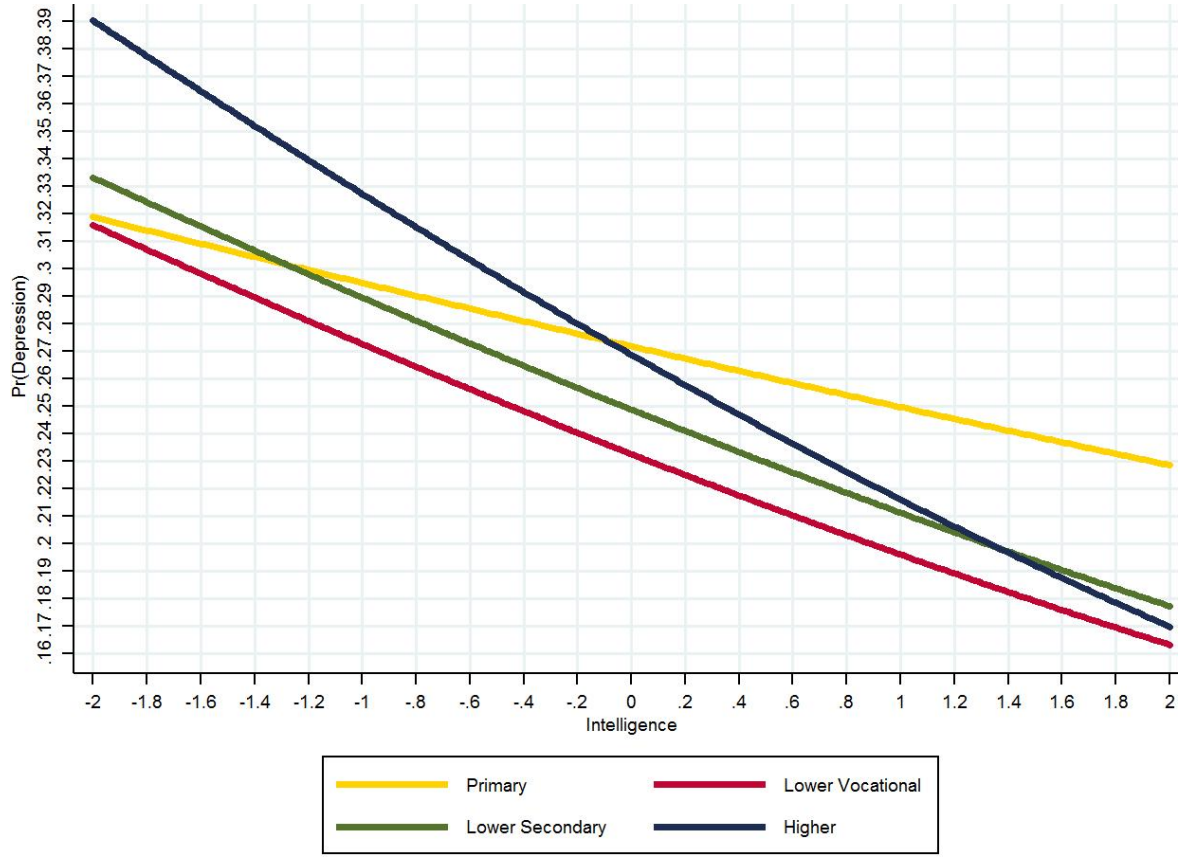


Figure 5: Probability of heartburn by education and intelligence

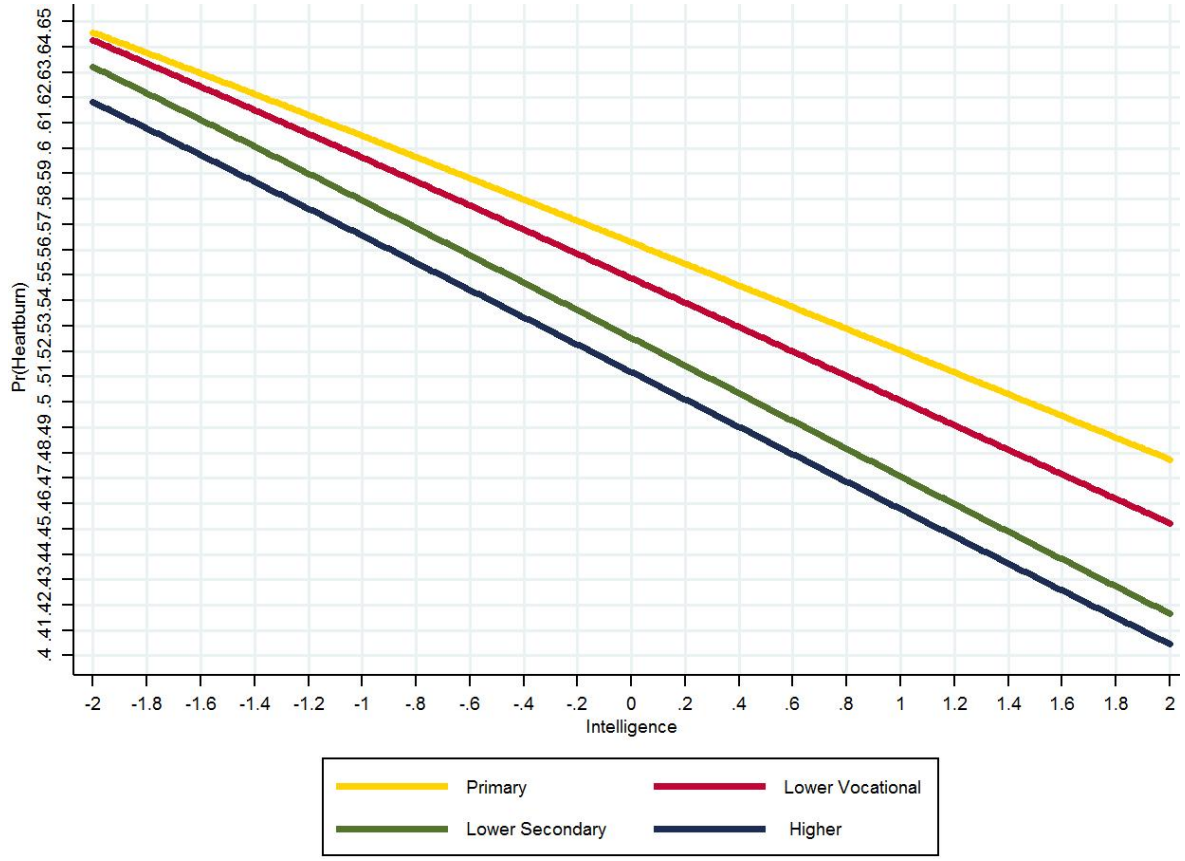


Figure 6: Probability of hypertension by education and intelligence

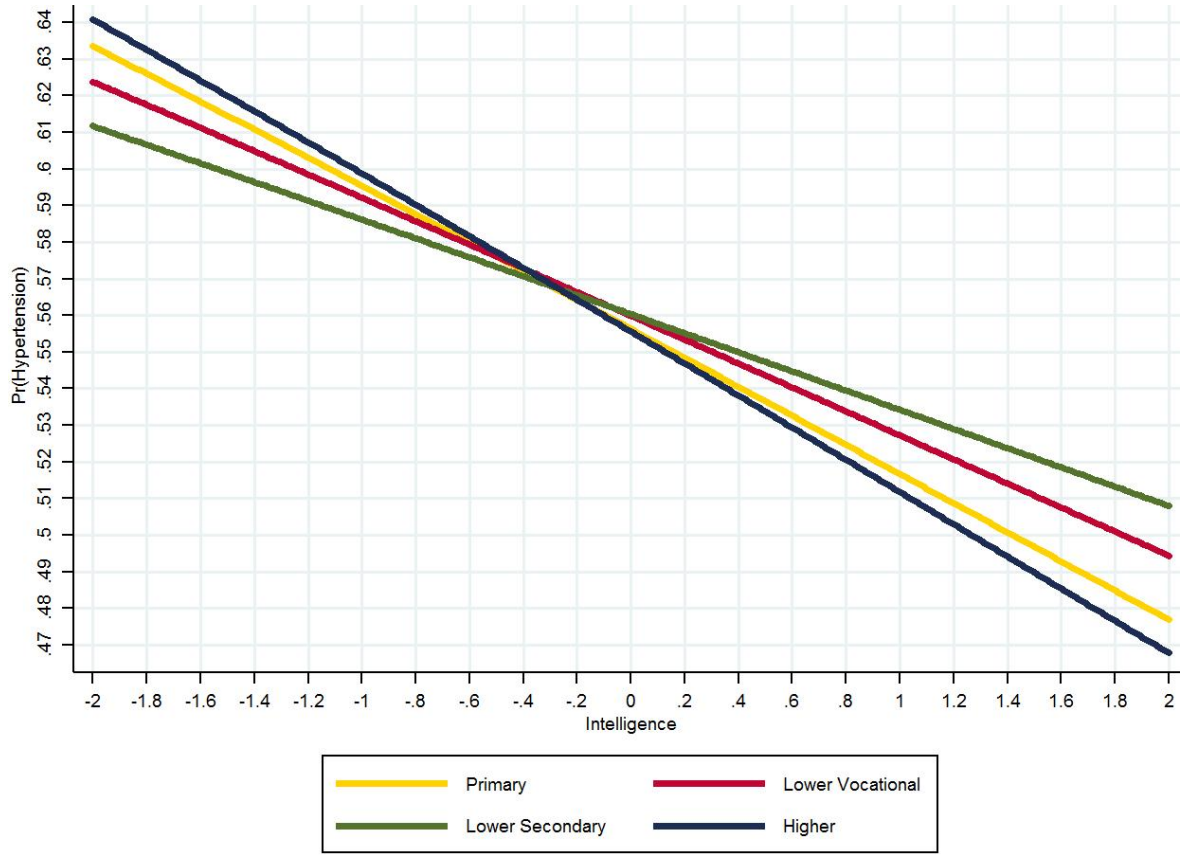


Figure 7: Probability of cardiac diseases by education and intelligence

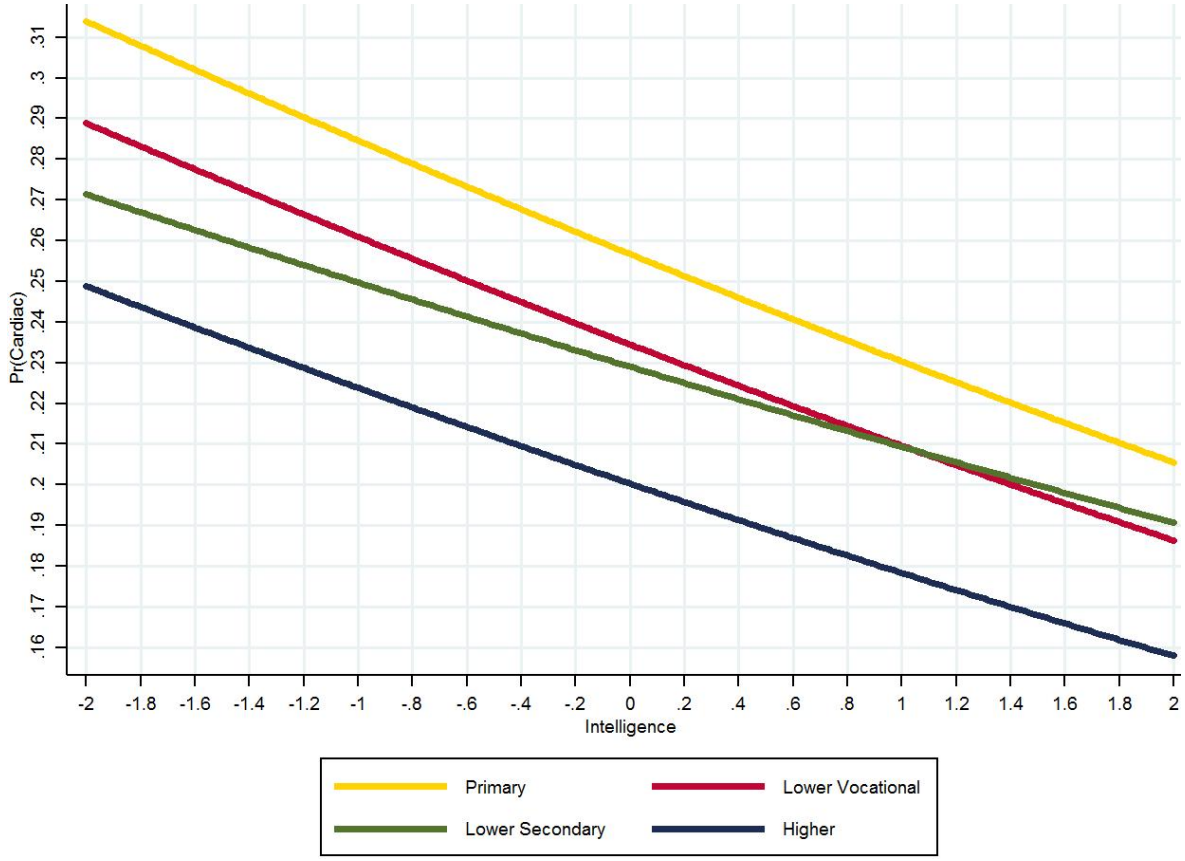
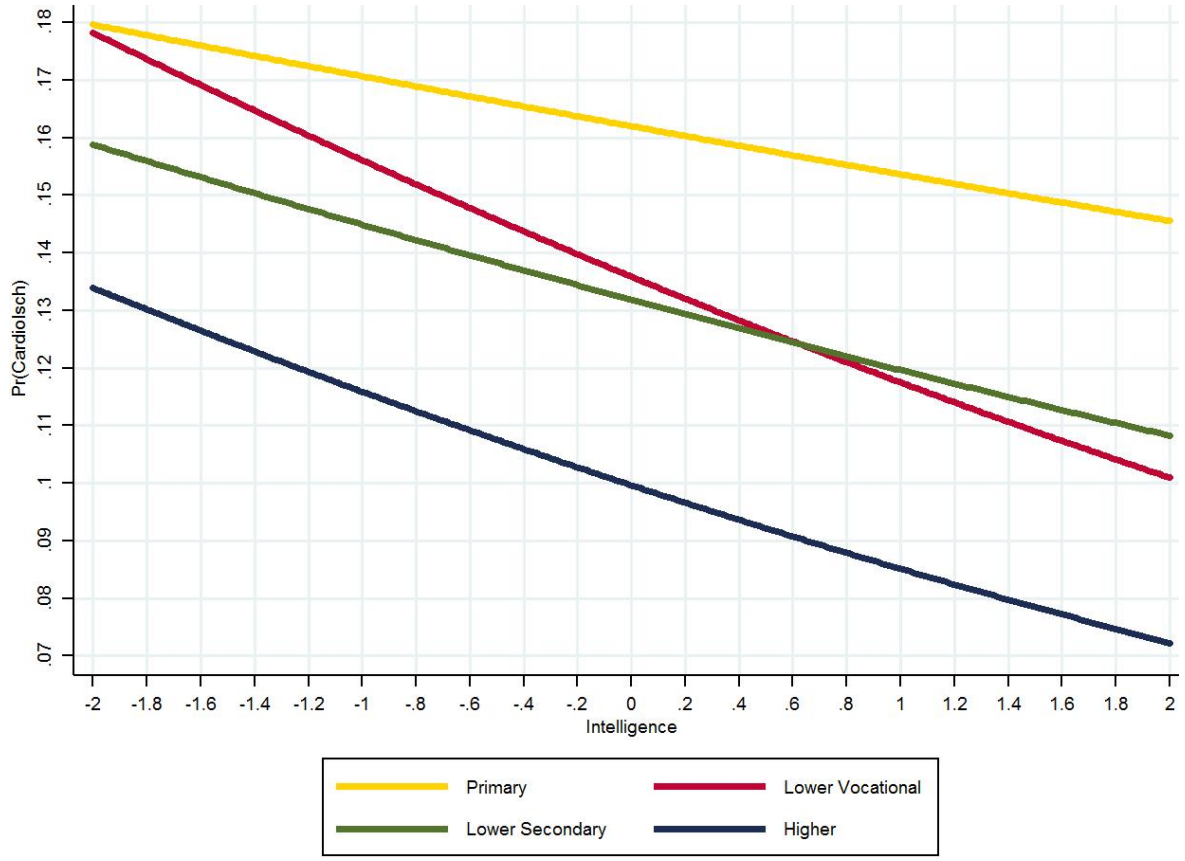


Figure 8: Probability of ischemic heart diseases by education and intelligence



3.4.2 Testing the differences in the predicted probabilities between different education levels

In this paragraph, the 95% and 90% confidence intervals (see equations (11) and (12)) of the difference $\Lambda_{d,h}^{(m)}$ in the predicted probabilities between different education levels are computed, at the mean of the observed covariates, where $d, h = 1, 2, 3, 4$ (primary, lower vocational, lower secondary, higher), $d < h$, and $\theta \in [-2; 2]$. In Figures 9-16, the orange and blue dots represent the 90% and 95% confidence intervals respectively. These allow to test the significance of the educational gain (namely, the decrease in the predicted probabilities from a certain level of education to a higher one) or loss of an average individual. Obviously, it is possible to plot the confidence intervals of the difference in the predicted probabilities at other levels of the observed covariates. This is valuable in order to understand the educational gain of a specific individual (namely, not an average individual) improving from d to h. If the confidence interval includes the value 0 (dashed line in the graphs), then the difference in the predicted probability is not significant at the respective level of θ.

According to the 95% confidence intervals, at some values of the latent intelligence the difference between primary and higher, $\Lambda_{1,4}^{(m)}$, is significantly different from 0 for the majority of the chronic diseases. Exceptions include depression and hypertension. As regards hyperlipidemia, diabetes, COPD, and cardiac diseases (including ischemic heart disease), $\Lambda_{1,4}^{(m)}$ is not significantly different from 0 at the highest or lowest values of the intelligence θ, but significant at intermediate values.

Most diseases are characterized by a significant difference between lower vocational and higher, $\Lambda_{2,4}^{(m)}$, at the 95% confidence level, usually from the lowest ($\theta_{min} = -2$) to medium-high values of θ. Exceptions include differences in hypertension and heartburn predicted probabilities, since the former is never significant and the latter is significant for low levels of θ only. It should be noted that the difference in the probabilities of depression $\Lambda_{2,4}^{(m)}$ is negative, which implies that improving the education level leads

to a significantly higher probability of the disease.

The primary and lower secondary differences $\Lambda_{1,3}^{(m)}$ of hyperlipidemia, COPD, depression, heartburn, and cardiac diseases (including ischemic heart disease) are significant from medium to high values of θ. This implies a significant educational gain for average individuals of medium to high intelligence, improving the education level from primary to lower secondary.

Computing and testing the differences in the predicted probabilities between two consecutive education levels $\Lambda_{1,2}^{(m)}$, $\Lambda_{2,3}^{(m)}$, and $\Lambda_{3,4}^{(m)}$ (primary vs lower vocational, lower vocational vs lower secondary, and lower secondary vs higher respectively) might be the most valuable results. Indeed, individuals commonly choose either to stop at a certain level of education or to continue to the consecutive one. Therefore, testing the significance of primary and lower vocational, lower vocational and lower secondary, and lower secondary and higher differences is useful in order to understand whether there is any significant educational gain by obtaining the consecutive level of education. Overall, the difference between primary and lower vocational, $\Lambda_{1,2}^{(m)}$, is significant for COPD, depression, and ischemic heart disease, and from medium to high levels of intelligence; the difference between lower vocational and lower secondary, $\Lambda_{2,3}^{(m)}$, is significant for depression and heartburn, and from medium to moderately high values of intelligence only; finally, the difference between lower secondary and higher, $\Lambda_{3,4}^{(m)}$, is significant for hyperlipidemia, diabetes, COPD, depression, and ischemic heart disease, from low to medium values of θ. Depending on the disease, and excluding depression, these results suggest that the average individual would eventually be significantly better off by obtaining the next education level in the following cases: if the level of intelligence is moderate to high and the level of education is low (primary); if the level of intelligence is around the average and the level of education is moderately low (lower vocational); if the level of intelligence is low to moderate and the level of education is moderately high (lower secondary). Overall, the null hypothesis of no significant educational gain (or loss) cannot be rejected at other levels of intelligence.

Figure 9: Educational gain in the probability of hyperlipidemia

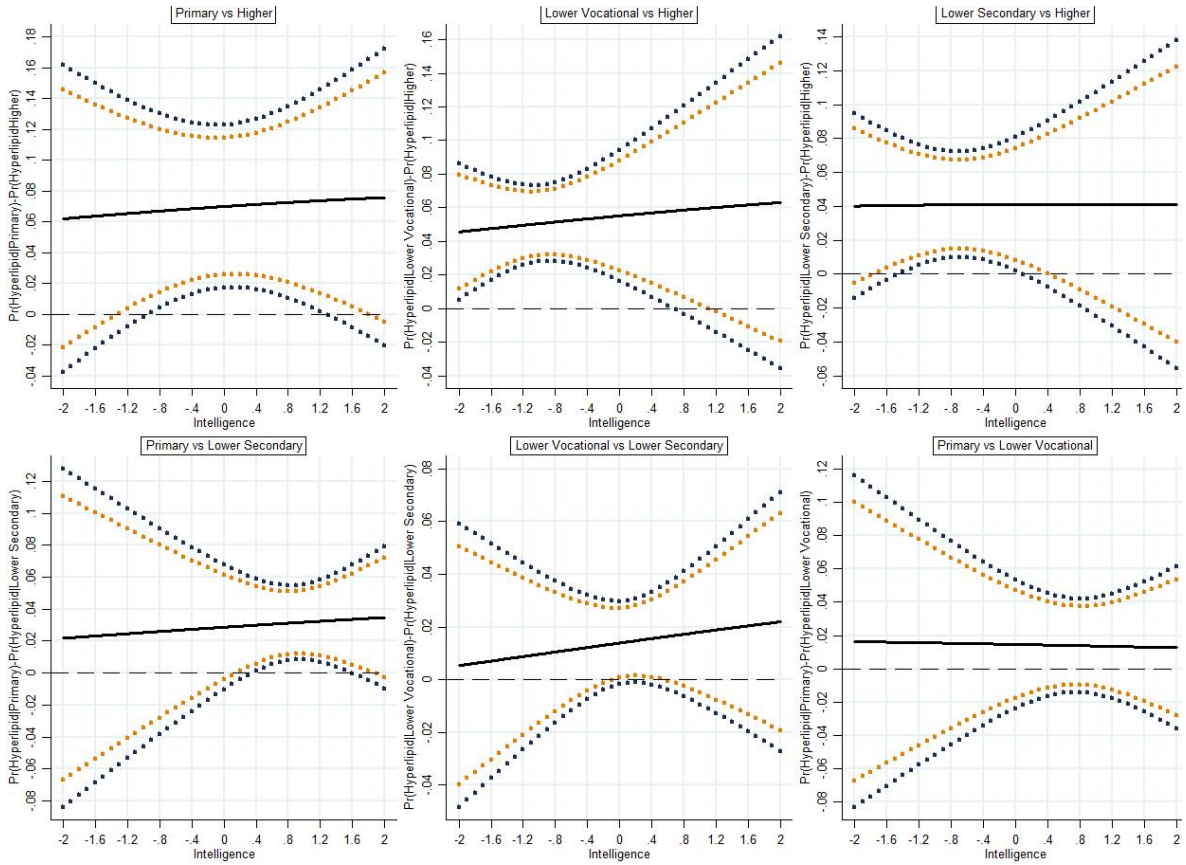


Figure 10: Educational gain in the probability of diabetes

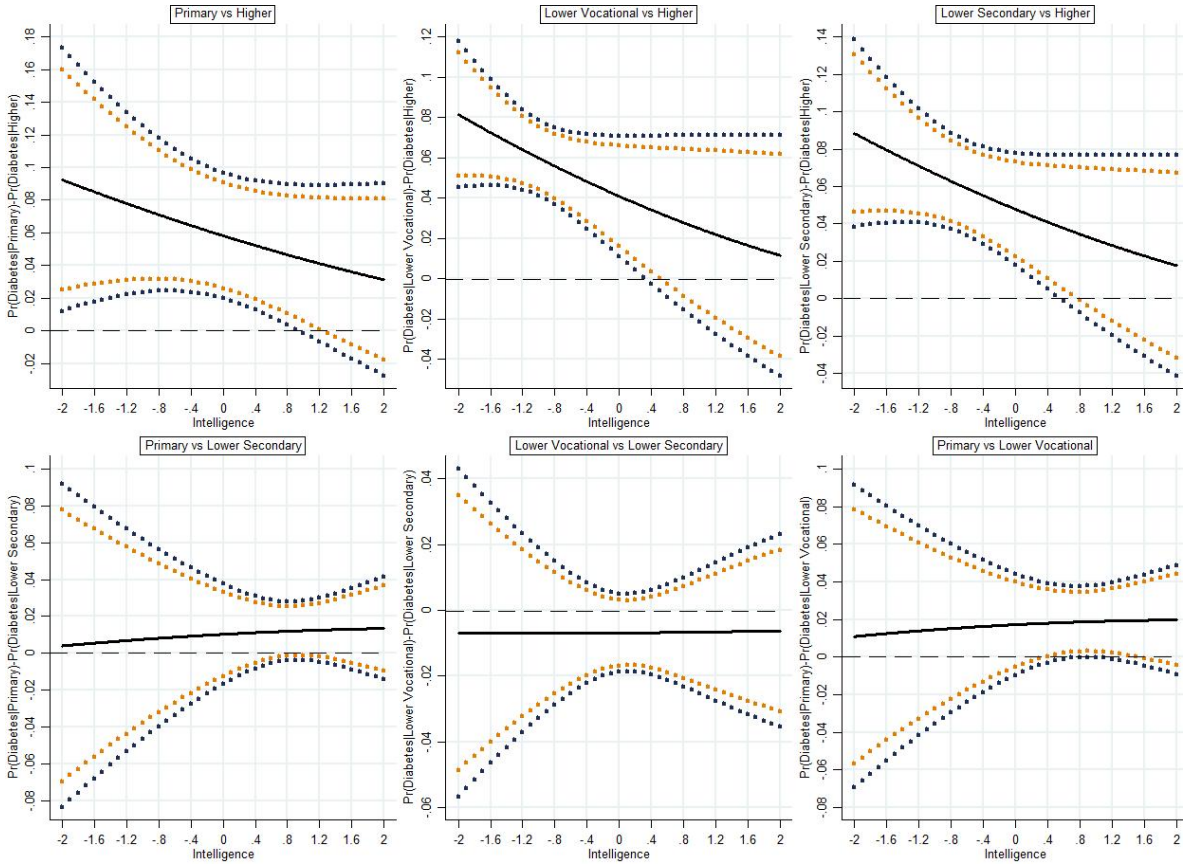


Figure 11: Educational gain in the probability of COPD

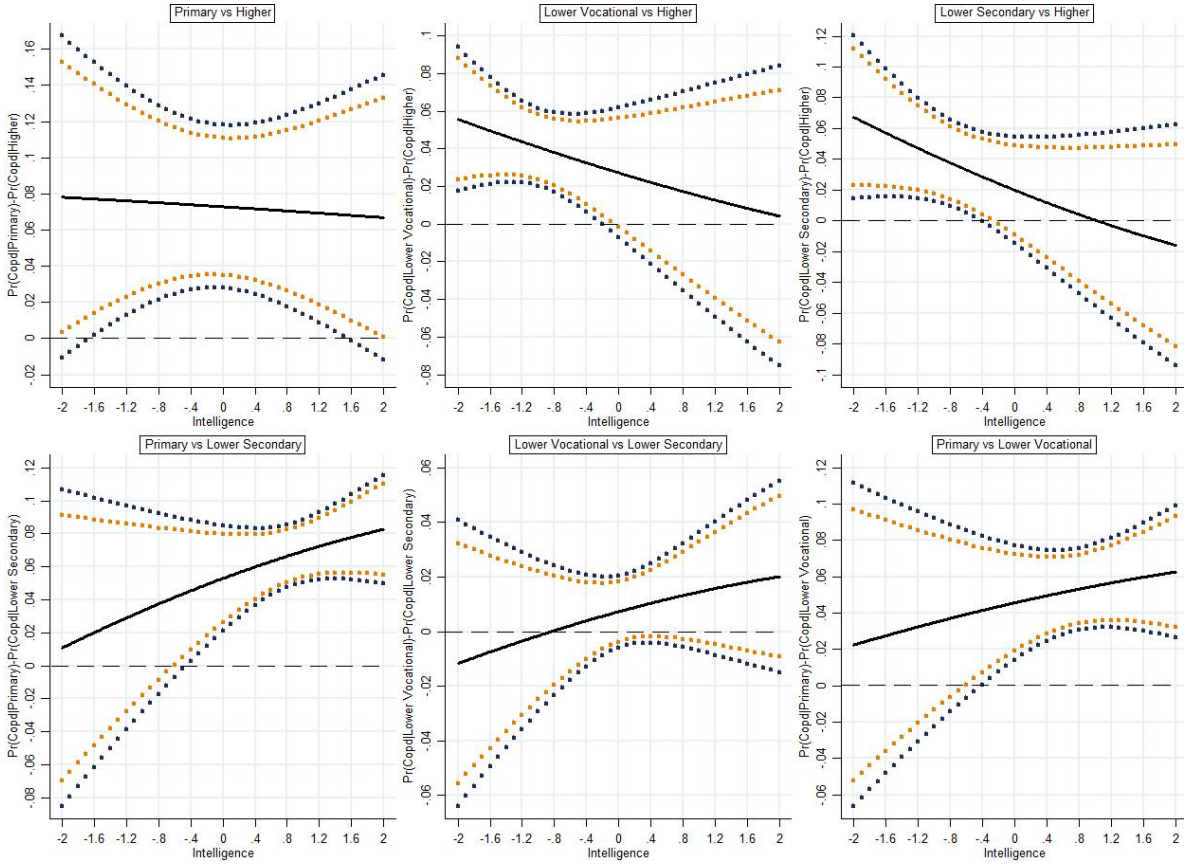


Figure 12: Educational gain in the probability of depression

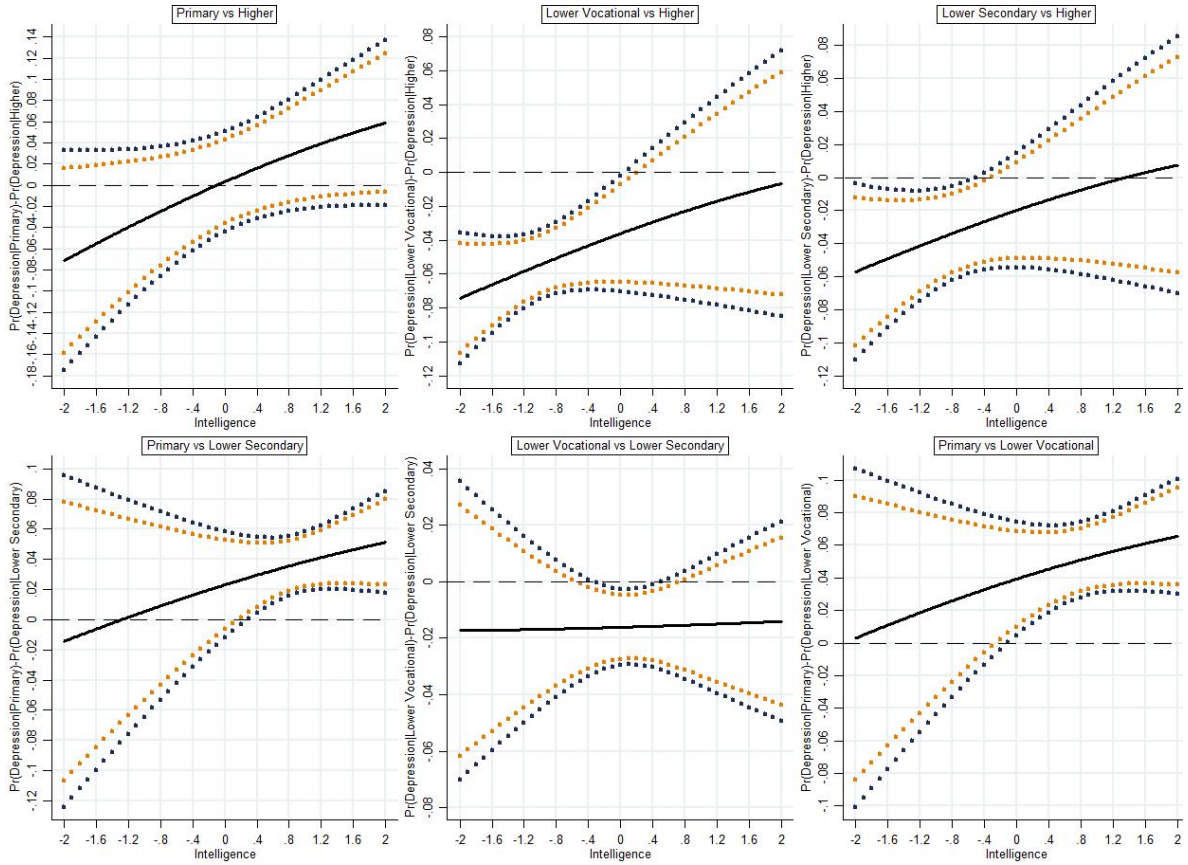


Figure 13: Educational gain in the probability of heartburn

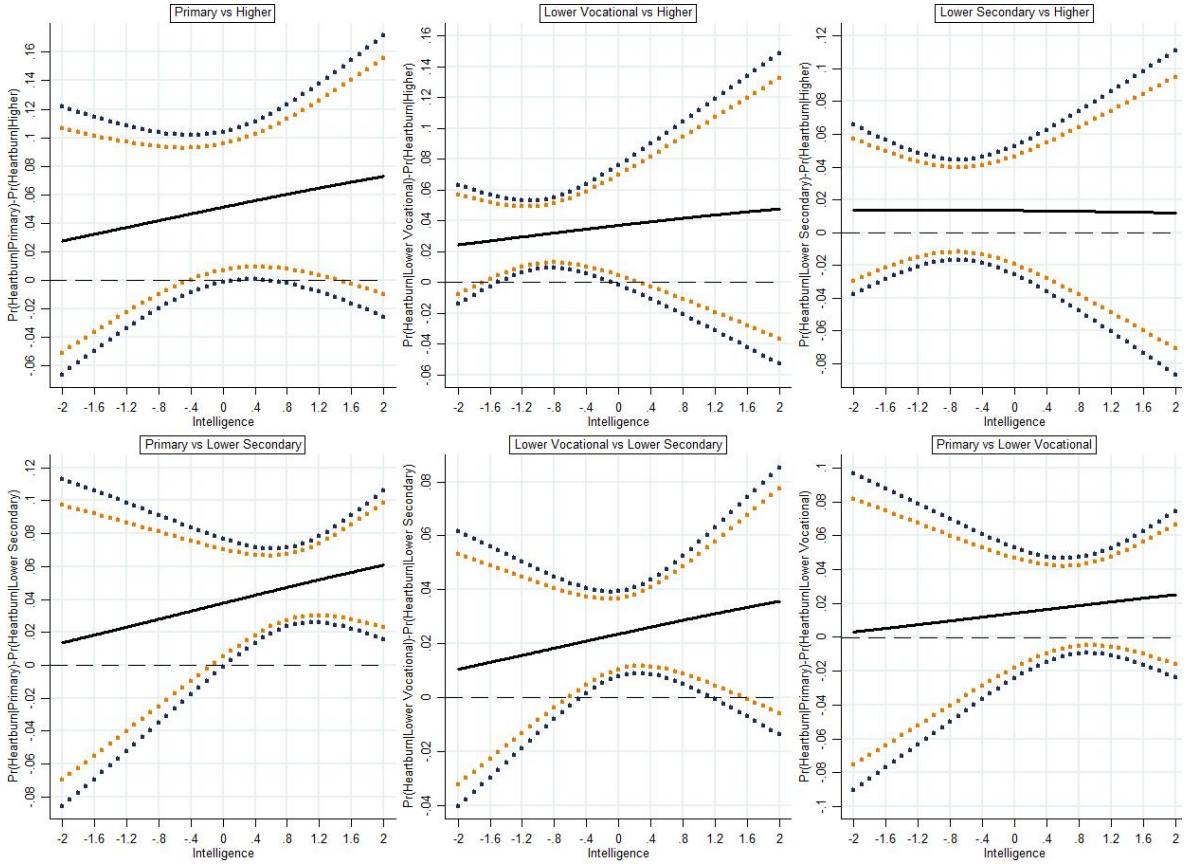


Figure 14: Educational gain in the probability of hypertension

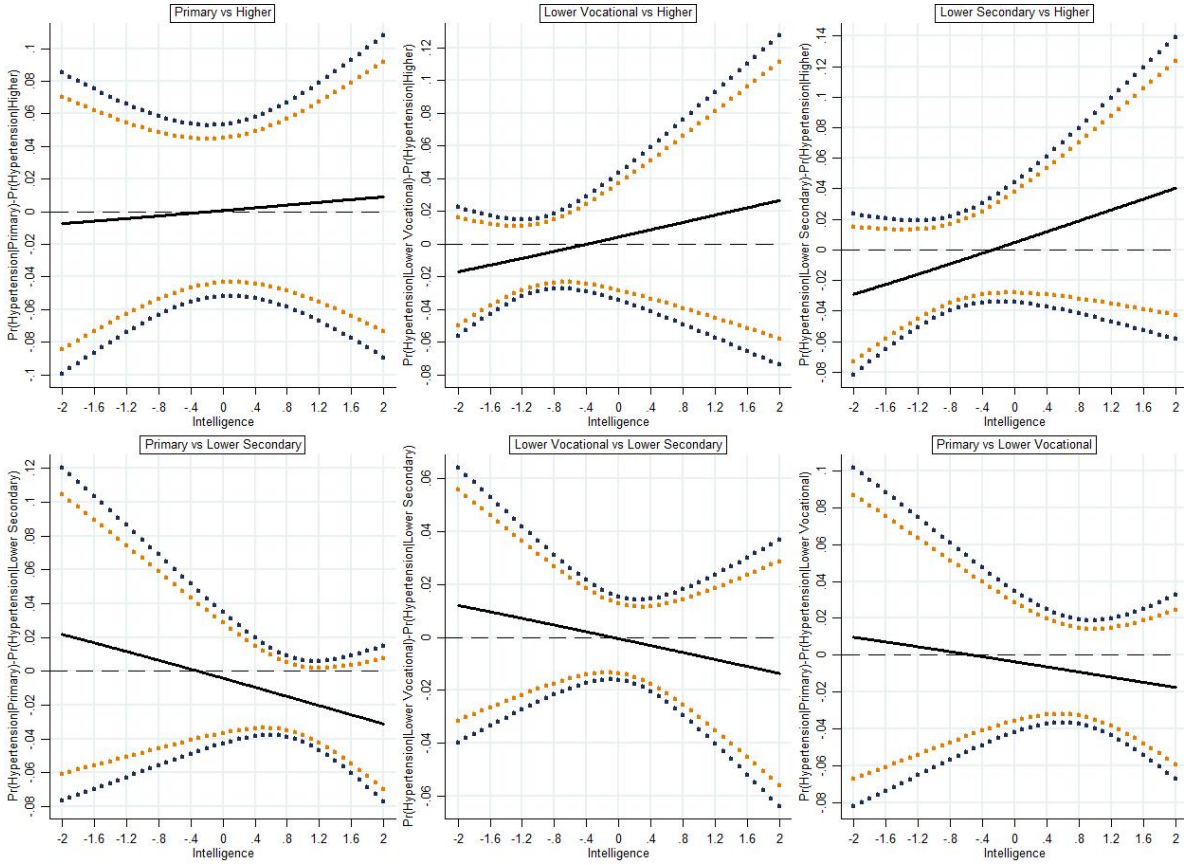


Figure 15: Educational gain in the probability of cardiac diseases

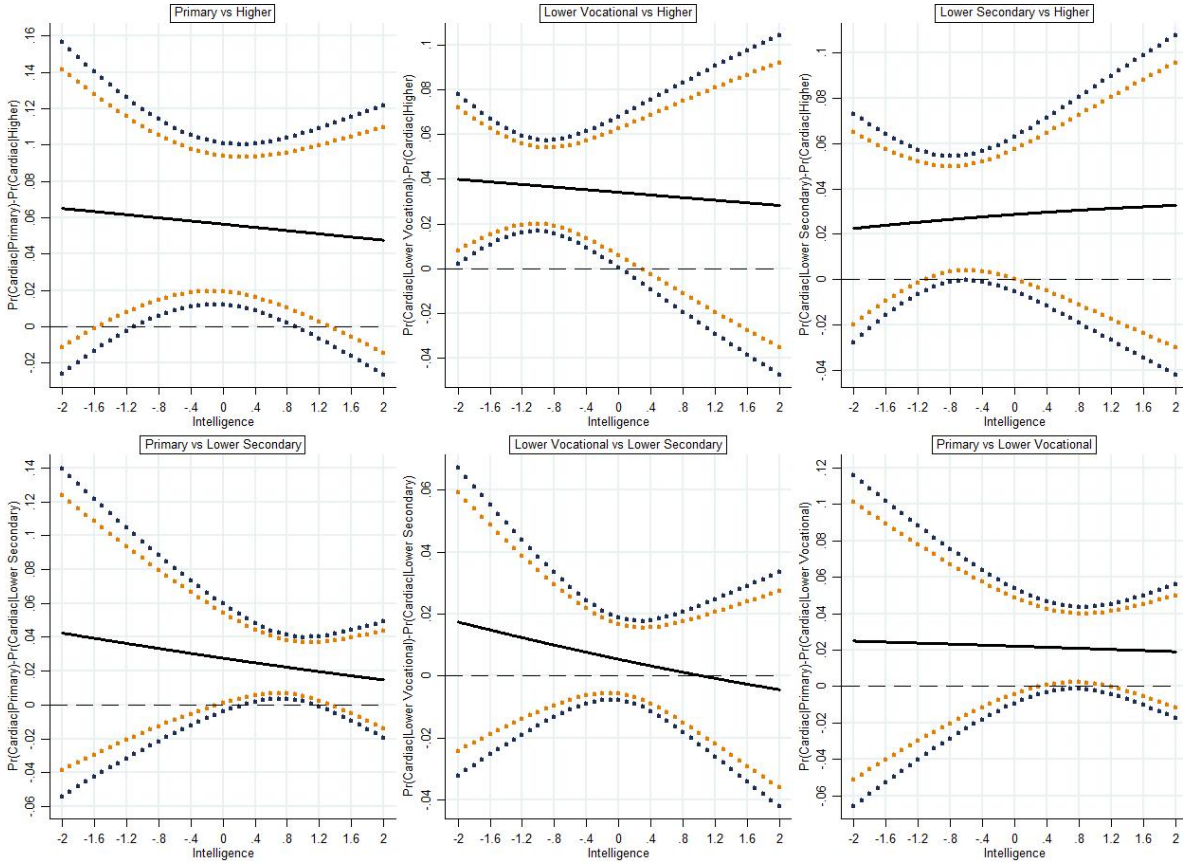
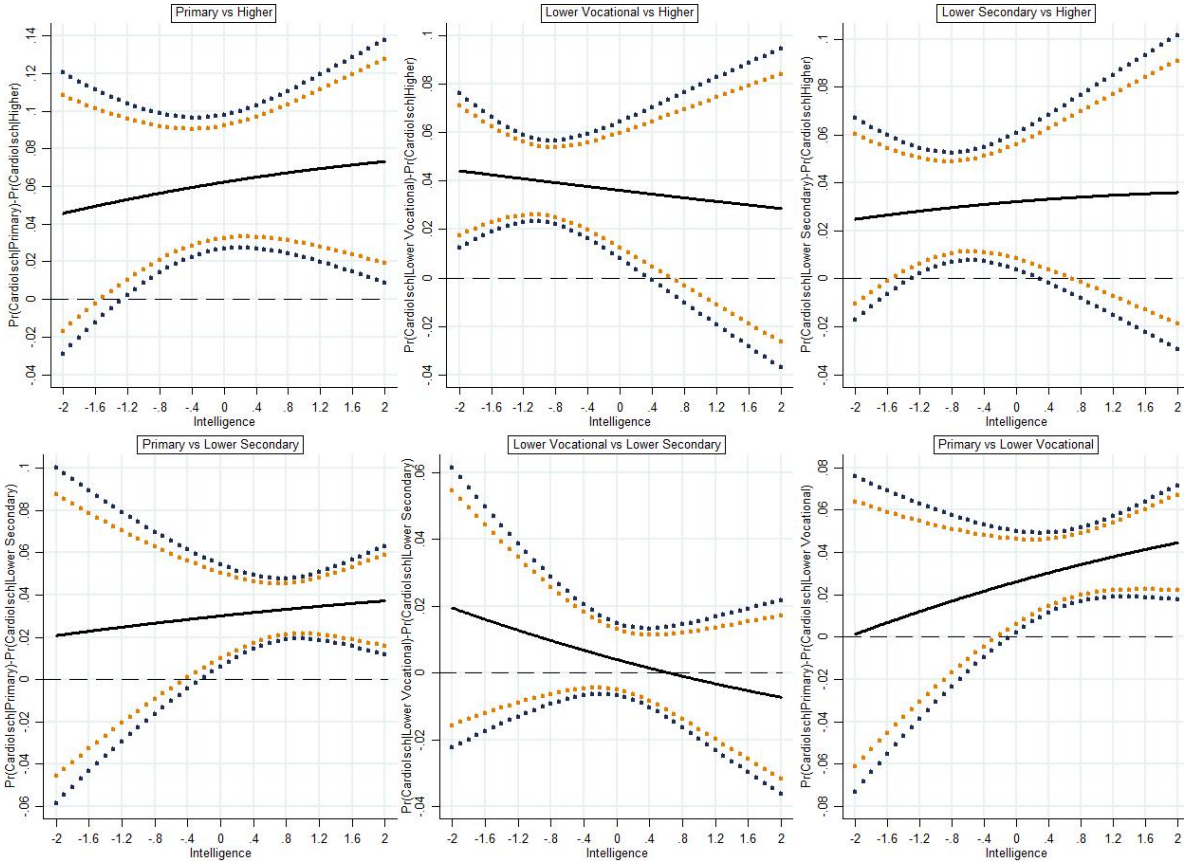
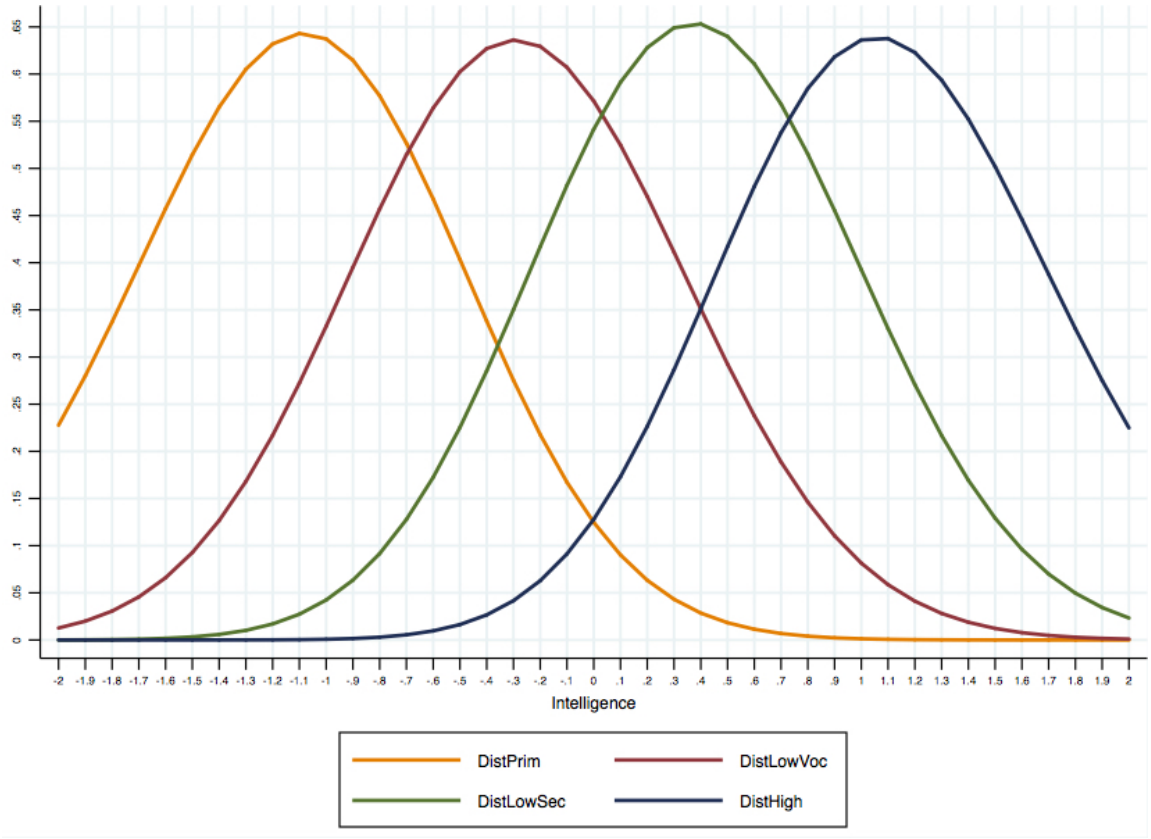


Figure 16: Educational gain in the probability of ischemic heart disease



3.4.3 Posterior distribution of intelligence

Figure 17: Distribution of intelligence given education and conditioning on $C = \bar{c}$



Following equation (14), the posterior distribution of the latent intelligence θ, given the level of education, for an imaginary population of average individuals (namely, with observed covariates $C = \bar{c}$) is represented in Figure 17. The expected value of the latent intelligence for each education level, and therefore for each subgroup of the population with primary, lower vocational, lower secondary, and higher education, is computed by means of equation (15). In particular, the expected values of the latent intelligence are $\bar{\theta}_1 = -1.122$, $\bar{\theta}_2 = -0.276$, $\bar{\theta}_3 = 0.400$, and $\bar{\theta}_4 = 1.117$ respectively.

Recall equation (16) that defines the percentage $\Gamma_{m,d}$ of average individuals with education level $d = 1, 2, 3$ (primary, lower vocational, and lower secondary) that, with confidence level 95%, would be less (or more) likely to use the medication m later in life, and thus to develop the respective disease, by going from education level d to $d + 1$. In order to compute $\Gamma_{m,d}$, it is necessary to observe, for each medication m, the

Table 16: Interval of intelligence for which an average individual would significantly gain (or lose) from choosing the consecutive education level.

Diseases or conditions	Primary	Lower vocational	Lower secondary
Hyperlipidemia			$[-1.4; 0]$
Diabetes			$[-2; 0.5]$
COPD (and bronchitis, asthma)	$[-0.4; 2]$		$[-2; -0.5]$
Depression (and anxiety)	$[-0.1; 2]$	$[-0.3; 0.5]$	$[-2; -0.5]$
Heartburn		$[-0.4; 1.1]$	
Hypertension			
Cardiac diseases			
Ischemic heart disease	$[0; 2]$		$[-1.3; 0.2]$

Table 17: Percentage of population of average individuals with certain education that would significantly gain (or lose) from choosing the next education level.

Diseases or conditions	Primary	Lower vocational	Lower secondary
Hyperlipidemia			$^{(-)}25.8\%$
Diabetes			$^{(-)}57.1\%$
COPD (and bronchitis, asthma)	$^{(-)}12.2\%$		$^{(-)}6.7\%$
Depression (and anxiety)	$^{(-)}4.7\%$	$^{(+)}40.3\%$	$^{(+)}6.7\%$
Heartburn		$^{(-)}56.0\%$	
Hypertension			
Cardiac diseases			
Ischemic heart disease	$^{(-)}3.3\%$		$^{(-)}37.5\%$

$^{(+)}$ more likely to get a disease improving educ.

$^{(-)}$ less likely to get a disease improving educ.

lowest ($\theta_{1,m,d}$) and highest ($\theta_{2,m,d}$) values of the latent intelligence $\theta \in [-2; 2]$ between which the differences in the probabilities between consecutive education levels, $\Lambda^{(m)}_{d,d+1}$, are significantly different from 0, at the 95% confidence level. From the 95% confidence intervals of $\Lambda^{(m)}_{d,d+1}$ (Figures 9-16), the intervals $[\theta_{1,m,d}; \theta_{2,m,d}]$ are observed and shown in Table 16, for each education level $d = 1, 2, 3$ and medication use m. Finally, the percentages of the population with primary, lower vocational, and lower secondary education, $\Gamma_{m,d}$, are computed and presented in Table 17.

As already mentioned in section 3.4.2, it is worth testing the significance of $\Lambda^{(m)}_{1,2}$, $\Lambda^{(m)}_{2,3}$, and $\Lambda^{(m)}_{3,4}$ in order to understand if there is any significant educational gain in the predicted probabilities, improving from a certain education level to the consecutive one.

Indeed, individuals commonly choose either to stop at a certain education level or to continue to the next one. Table 17 shows even more useful results, taking advantage of the posterior distribution of the latent intelligence. Overall, only for few diseases and education levels there is a notable percentage of average individuals significantly better off continuing to the consecutive education level: 12.2% of the group with primary education improving to lower vocational (COPD); 56.0% of the group with lower vocational education improving to lower secondary (heartburn); 57.1%, 37.5%, and 25.8% of the group with lower secondary education improving to higher (diabetes, ischemic heart disease, and hyperlipidemia respectively). For other diseases, the percentages appear to be negligible or represent individuals who would be worse off improving the education level.

If the real population is sufficiently represented by the imaginary population of average individuals, then incrementing one education level further may lead to only a percentage of individuals significantly reducing the probability of getting some diseases, according to their level of intelligence and education. However, this assumption may be difficult to justify. In this sense, a strategy for dealing with the issue is testing the significance of the difference in the predicted probabilities between education levels at other different values of the observed covariates. This allows to quantify the educational gain and consequently compute the percentage of individuals with those characteristics being less likely to get the disease. Comparing the educational gain for the population of average individuals with a more general measure of causal effect of education (namely, the treatment effect) and verifying whether the results are coherent is also possible. This is analysed and discussed is section 3.4.4.

3.4.4 ATE, TT, and TU

Following equation (19), the average treatment effects from choosing lower vocational, lower secondary, and higher education level rather than the previous one are computed, for each disease. The results are shown in Table 18. Similarly, the treatment

Table 18: Average treatment effect in percentage points improving to lower vocational, lower secondary, and higher education from the prior level.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	-1.77	-1.52	-4.76^{**}
	(-0.83)	(-1.85)	(-2.81)
Diabetes	-1.22	$+0.88$	-4.34^{***}
	(-0.91)	(1.39)	(-3.59)
COPD (and bronchitis, asthma)	-4.23^{**}	-0.60	-1.79
	(-2.73)	(-0.86)	(-1.14)
Depression (and anxiety)	-3.44	$+1.63^{*}$	$+2.24$
	(-1.86)	(2.26)	(1.39)
Heartburn	-1.19	-2.30^{***}	-1.32
	(-0.69)	(-3.29)	(-0.83)
Hypertension	$+0.50$	-0.00	-1.30
	(0.28)	(-0.00)	(-0.74)
Cardiac diseases	-2.21	-0.50	-2.78
	(-1.27)	(-0.73)	(-1.95)
Ischemic heart disease	-2.59	-0.43	-2.75^{*}
	(-1.95)	(-0.78)	(-2.55)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 19: Treatment effect on the treated in percentage points improving to lower vocational, lower secondary, and higher education from the prior level.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	-1.85	-1.43	-5.29^{***}
	(-0.87)	(-1.78)	(-3.47)
Diabetes	-1.38	$+0.64$	-3.91^{***}
	(-0.88)	(1.04)	(-3.62)
COPD (and bronchitis, asthma)	-4.33^{*}	-0.72	-1.70
	(-2.32)	(-1.05)	(-1.33)
Depression (and anxiety)	-3.63	$+1.56^{*}$	$+2.38$
	(-1.95)	(2.26)	(1.64)
Heartburn	-1.22	-2.52^{**}	-1.20
	(-0.56)	(-3.15)	(-0.78)
Hypertension	$+0.31$	-0.23	-2.21
	(0.14)	(-0.29)	(-1.46)
Cardiac diseases	-2.38	-0.49	-2.49
	(-1.30)	(-0.72)	(-1.95)
Ischemic heart disease	-2.71	-0.43	-2.11^{*}
	(-1.73)	(-0.78)	(-2.10)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 20: Treatment effect on the untreated in percentage points improving to lower vocational, lower secondary, and higher education from the prior level.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	-1.57	-1.61	-4.51^{*}
	(-0.76)	(-1.85)	(-2.50)
Diabetes	-0.80	$+1.10$	-4.55^{***}
	(-0.52)	(1.56)	(-3.43)
COPD (and bronchitis, asthma)	-3.98^{*}	-0.49	-1.83
	(-2.19)	(-0.64)	(-1.21)
Depression (and anxiety)	-2.94	$+1.69^{*}$	$+2.18$
	(-1.62)	(2.18)	(1.27)
Heartburn	-1.11	-2.10^{*}	-1.38
	(-0.53)	(-2.42)	(-0.77)
Hypertension	$+1.01$	$+0.20$	-0.86
	(0.49)	(0.23)	(-0.48)
Cardiac diseases	-1.76	-0.51	-2.91
	(-1.00)	(-0.68)	(-1.92)
Ischemic heart disease	-2.26	-0.43	-3.06^{**}
	(-1.47)	(-0.69)	(-2.61)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

effects on the treated and untreated individuals are also computed and shown in Table 19 and Table 20 respectively. The results are in percentage points.

It can be seen that the primary to lower vocational treatment effects are significant at the 95% confidence level (or above) for COPD only. In particular, by choosing to continue from primary to lower vocational, the treated individuals have significantly reduced the predicted probability of getting COPD (or chronic bronchitis, asthma) by an average of 4.33 percentage points, and the untreated would have significantly reduced it by an average of 3.98 percentage points. No other lower vocational treatment effect is significant at the 95% confidence level, even though the test statistics of the average treatment effect and the treatment effect on the treated for ischemic heart disease and depression respectively are close to the threshold of 1.96. Similarly, the lower vocational to lower secondary treatment effects are significant for depression and heartburn. Thus, similarly to what obtained in sections 3.4.2 and 3.4.3, continuing to lower secondary education rather than ending at lower vocational leads to an average sig-

nificant increase in the predicted probability of getting depression (or anxiety) equal to around 1.6 percentage points, both for the treated and the untreated. Finally, the lower secondary to higher treatment effects are significant and negative at the 95% confidence level (or above) for hyperlipidemia, diabetes, and ischemic heart disease. Moreover, the test statistics of the lower secondary to higher treatment effects for cardiac diseases are close to 1.96.

It may be worth comparing the treatment effects with the percentages in Table 17. Almost always, the treatment effects are significant if the respective percentage of the population of average individuals is greater than 0. That is, if a not null percentage of the imaginary average group with a certain education level would be better or worse off continuing to the consecutive level, then the respective treatment effect is significant, and vice versa. For instance, 57.1% of the population of average individuals, with lower secondary education, would significantly decrease the probability of diabetes improving to higher education, and the respective lower secondary to higher treatment effects are strongly significant. Exceptions include ischemic heart disease and depression from primary to lower vocational, and COPD and depression from lower secondary to higher. Indeed, the respective treatment effects are not significantly different from zero. However, as already mentioned in section 3.4.3, the percentages from primary to lower vocational for ischemic heart disease and depression, and from lower secondary to higher for COPD and depression appear to be negligible with respect to the other not null percentages, being equal to only 3.3%, 4.7%, 6.7%, and 6.7% respectively.

Overall, the results are coherent to what is obtained in sections 3.4.2 and 3.4.3. Indeed, almost all the treatment effects are significant if the difference in probabilities between the two respective consecutive education levels are significantly different from zero, at a certain interval of the latent intelligence. This suggests that intelligence may play a fundamental role in determining the significance of the educational gain not only for the average (imaginary) individual, but also for the individuals in the dataset. If

this hypothesis is true, then the significance of the causal effect of education on the occurrence of chronic diseases in old age depends not only on the considered disease and education level, but also on the level of intelligence. In other words, given that few treatment effects are observed to be significant and that the respective percentages in Table 17 are never higher than 57.1%, the causal effect of improving from one education level to the consecutive one may be hardly ever significant.

Chapter 4

Conclusions

This thesis evaluates the causal effect of education on health outcomes later in life, through a structural model that controls for the interdependence among education choice, observed and unobserved confounders, and occurrence of chronic diseases in old age. The occurrence of chronic diseases is identified from medication use. The three-part structural model is composed of a sequential probit model for the educational attainment, a measurement system using IQ-tests to identify latent intelligence, and a probit model for medication use. The causal effect of education is mainly analysed with respect to improving the education from a certain level to the consecutive one.

Overall, the results offer little evidence of significant causal effect of education. After controlling for latent intelligence and observed confounders, education still negatively affects the occurrence of few chronic diseases in old age. Both for the treated and the untreated individuals in the dataset, the probabilities of *(i)* COPD *(ii)* heartburn *(iii)* hyperlipidemia, diabetes, ischemic heart disease, significantly decrease by improving the education level from *(i)* primary to lower vocational *(ii)* lower vocational to lower secondary *(iii)* lower secondary to higher respectively. The probability of depression significantly raises by improving the education level, from lower vocational to lower secondary.

The differences in the probabilities of an average individual being affected by the same diseases (COPD, depression, heartburn, hyperlipidemia, diabetes, and ischemic heart disease), are significant for the same consecutive education levels as for the treatment effects. However, these differences in the predictions are significant at some levels of intelligence only. This allows to compute the percentage of the population of average individuals that would be less or more likely to develop a certain disease in old age by improving the education level. The highest percentages of each education level are estimated for COPD (12.2% of the population of average individuals with primary education would be significantly better off improving to lower vocational), heartburn (56.0% of the population of average individuals with lower vocational education would be significantly better off improving to lower secondary), and diabetes (57.1% of the population of average individuals with lower secondary education would be significantly better off improving to higher). Moreover, except for the lowest ones, the percentages are generally different from zero if the treatment effects are significant, and vice versa. This suggests that the level of intelligence may play a fundamental role in determining whether or not improving the education level may significantly decrease the probability of getting a certain disease, not only for the average (imaginary) individual but also for the individuals in the dataset. In particular, according to the confidence intervals of the difference in predicted probabilities at different values of the latent intelligence, the average individual would be significantly better off by obtaining the consecutive education level in the following cases: if the level of intelligence is moderate to high and the education level is low (primary), if the level of intelligence is around the average and the education level is moderately low (lower vocational), if the level of intelligence is low to moderate and the education level is moderately high (lower secondary).[1]

The importance of using the structural model approach to avoid overestimating the causal effect of education is suggested by the distribution of IQ-tests scores and ob-

[1]However, it should be stressed that if for a certain disease or level of intelligence the difference in the predictions is not significant, then it does not mean that the null hypothesis of no causal effect of education is true, but only that cannot be rejected.

served confounders, by education level. On the one hand, individuals who have better observed characteristics, in terms of decreasing the chances of developing a chronic disease in old age, are those who have higher education levels. For instance, according to the descriptive statistics, people who have fathers whose job is professional are more likely to be higher educated; at the same time, having a wealthier family may contribute in decreasing the probability of developing chronic conditions later in life, through, for example, a more balanced diet over the lifespan. On the other hand, the IQ-tests scores increase with education, indicating that the latent intelligence (which is identified from the IQ-tests) may significantly contribute to obtaining higher education levels. Also, the importance of using the structural model approach comes from comparing the estimated treatment effects with the descriptive statistics on the occurrence of chronic diseases by education level. Overall, the average treatment effects are lower than what is suggested by the descriptive statistics. For instance, there is a decrease in the percentage of hyperlipidemia medication use of -8.4 percentage points from lower secondary to higher education, while the respective average treatment effect is equal to -4.76 percentage points. Moreover, the percentage of individuals who suffer from depression (or anxiety) decreases with education, but the respective average treatment effects are positive from lower vocational to lower secondary and from lower secondary to higher education. Finally, the treatment effects from a (non-structural) medication use probit model are compared with the structural model ones. This further justifies the importance of the structural model approach. Indeed, the majority of treatment effects from the non-structural model are significantly overestimated with respect to the ones from the structural model.

The contribution of this thesis is therefore threefold. Firstly, according to the treatment effects, the causal effect of improving the education from a certain level to the consecutive one is significant for few diseases and education levels. Secondly, at some intervals of the latent intelligence, and thus for a large part of the population, the same causal effect of education may be not significantly different from zero. Thirdly, not con-

trolling for the unobserved endogeneity would lead to overestimating the causal effect of education. In other words, after controlling for the endogeneity the significance of the causal effect of education depends not only on the considered disease and education level, but also on the level of intelligence. The causal effect of improving from a certain education level to the consecutive one, on the occurrence of chronic diseases later in life, is therefore hardly ever significant.

Appendix A

A simple non-structural probit model for each chronic disease is estimated (Tables A4-A11). Moreover, the treatment effects from primary to lower vocational, from lower vocational to lower secondary, and from lower secondary to higher education are computed (Tables A1, A2, and A3). These are compared to the ones from the structural model (Tables 18, 19, and 20). It can be seen that the treatment effects for each education level are systematically lower than the ones from the structural model. In other words, not including an educational attainment model and not controlling for the latent intelligence that affects both the education choice and the occurrence of diseases, leads to systematically overestimating the educational gain. Moreover, it is tested whether the treatment effects based on the simple probit model are significantly different, at the 95% confidence level, from the ones based on the structural model.[2] It can be seen that the lower secondary treatment effects lie outside the 95% confidence intervals of the respective ones based on the structural model, except for ischemic heart disease. Lower vocational average treatment effect for diabetes and all the higher education treatment effects for depression, heartburn, and hypertension are also significantly different. Finally, for some diseases such as depression, the sign of the estimated treatment effects changes from positive (structural model) to negative (simple probit model). Thus, evidence suggests that confounders, such as latent intelligence, play an important role in shaping the association between education and health.

[2]Tables 18, 19, 20 report the point estimates and the t statistics in parentheses. Thus, the standard errors and consequently the 95% confidence intervals of the treatment effects based on the structural model can be computed.

Table A1: Average treatment effect improving to lower vocational, lower secondary, and higher education from the prior level. Non-structural model.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	−4.34	−3.67*	−7.21
Diabetes	−3.87*	−1.37*	−4.88
COPD (and bronchitis, asthma)	−5.81	−3.05*	−2.50
Depression (and anxiety)	−5.01	−0.99*	−1.98*
Heartburn	−4.52	−5.78*	−5.25*
Hypertension	−2.90	−1.94*	−4.92*
Cardiac diseases	−4.58	−2.02*	−4.42
Ischemic heart disease	−3.08	−1.48	−3.87

* outside 95% c.i. ATE (Table 18)

Table A2: Treatment effect on the treated improving to lower vocational, lower secondary, and higher education from the prior level. Non-structural model.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	−4.41	−3.49*	−7.69
Diabetes	−4.03	−1.52*	−4.60
COPD (and bronchitis, asthma)	−5.89	−3.05*	−2.61
Depression (and anxiety)	−5.19	−0.97*	−1.65*
Heartburn	−4.55	−5.86*	−5.07*
Hypertension	−3.09	−2.14*	−5.61*
Cardiac diseases	−4.75	−2.00*	−4.11
Ischemic heart disease	−3.19	−1.47	−3.23

* outside 95% c.i. TT (Table 19)

Table A3: Treatment effect on the untreated improving to lower vocational, lower secondary, and higher education from the prior level. Non-structural model.

Diseases or conditions	Lower vocational	Lower secondary	Higher
Hyperlipidemia	−4.14	−3.83*	−6.98
Diabetes	−3.45	−1.23*	−5.01
COPD (and bronchitis, asthma)	−5.58	−3.04*	−2.44
Depression (and anxiety)	−4.54	−1.00*	−2.14*
Heartburn	−4.46	−5.72*	−5.33*
Hypertension	−2.38	−1.76*	−4.59*
Cardiac diseases	−4.13	−2.04*	−4.58
Ischemic heart disease	−2.78	−1.48	−4.18

* outside 95% c.i. TU (Table 20)

Table A4: Simple probit hyperlipidemia

Hyperlipidemia	Primary	Lower vocational	Lower secondary	Higher
Constant	0.046	0.056	−0.095	−0.147*
	(0.58)	(1.23)	(−1.93)	(−2.00)
BMI	0.006	0.021***	0.027***	0.027**
	(0.75)	(4.07)	(4.93)	(3.14)
Height	−0.011***	−0.013***	−0.017***	−0.018***
	(−3.76)	(−7.29)	(−9.03)	(−6.27)
Lower than fit health	0.041	0.016	0.057	−0.079
	(0.81)	(0.49)	(1.73)	(−1.61)
Lower than fit psych	−0.022	−0.035	−0.028	−0.003
	(−0.56)	(−1.22)	(−0.87)	(−0.06)
Reference category: White collar				
Professional	−0.040	−0.081	0.032	0.006
	(−0.54)	(−1.96)	(0.91)	(0.15)
Farm owner	0.098	−0.194***	−0.248**	0.180
	(0.82)	(−3.49)	(−2.78)	(1.26)
Skilled	0.100	0.028	0.017	0.052
	(1.95)	(0.98)	(0.54)	(0.81)
Unskilled	0.080	0.026	0.032	0.098
	(1.39)	(0.73)	(0.75)	(1.00)
Unknown	−0.020	0.084	0.148*	0.173
	(−0.26)	(1.64)	(2.55)	(1.83)
Reference category: Selected city				
Non-urban	−0.018	−0.049	−0.033	−0.064
	(−0.27)	(−1.27)	(−0.67)	(−0.84)
Urban	0.009	−0.039	−0.002	−0.105
	(0.12)	(−0.77)	(−0.03)	(−1.21)
Reference category: 1945				
1944	0.084	−0.020	0.028	−0.043
	(1.17)	(−0.46)	(0.59)	(−0.60)
1946	−0.002	−0.054	−0.123*	−0.165*
	(−0.03)	(−1.19)	(−2.40)	(−2.16)
1947	−0.025	−0.178***	−0.145**	−0.305***
	(−0.31)	(−3.60)	(−2.58)	(−3.65)
Famine exposure				
PN	−0.033	−0.110**	0.004	−0.105
	(−0.50)	(−2.80)	(0.10)	(−1.76)
T3	0.038	0.090*	0.020	−0.029
	(0.54)	(2.06)	(0.42)	(−0.42)
T2	0.014	−0.055	0.060	−0.082
	(0.20)	(−1.27)	(1.35)	(−1.24)
T1	−0.077	0.022	−0.006	−0.145
	(−0.92)	(0.42)	(−0.10)	(−1.75)
PC	−0.043	−0.070	0.037	−0.039
	(−0.68)	(−1.76)	(0.91)	(−0.67)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: Simple probit diabetes

Diabetes	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.506^{***}	-0.958^{***}	-1.026^{***}	-1.152^{***}
	(-5.80)	(-17.75)	(-17.51)	(-12.20)
BMI	0.039^{***}	0.049^{***}	0.062^{***}	0.069^{***}
	(4.60)	(8.28)	(9.65)	(6.56)
Height	0.002	-0.005^{*}	-0.012^{***}	-0.011^{**}
	(0.49)	(-2.41)	(-5.33)	(-2.94)
Lower than fit health	0.038	0.048	0.123^{**}	0.127^{*}
	(0.68)	(1.27)	(3.17)	(2.07)
Lower than fit psych	0.010	0.067^{*}	0.090^{*}	0.087
	(0.22)	(2.01)	(2.42)	(1.45)
Reference category: White collar				
Professional	-0.075	-0.039	-0.033	0.030
	(-0.91)	(-0.81)	(-0.78)	(0.59)
Farm owner	-0.236	-0.137^{*}	-0.109	-0.050
	(-1.66)	(-2.01)	(-0.99)	(-0.24)
Skilled	-0.081	0.006	0.042	0.074
	(-1.43)	(0.19)	(1.13)	(0.91)
Unskilled	-0.004	0.032	0.066	0.028
	(-0.06)	(0.77)	(1.30)	(0.22)
Unknown	-0.137	-0.016	0.078	0.203
	(-1.53)	(-0.26)	(1.15)	(1.73)
Reference category: Selected city				
Non-urban	-0.228^{**}	-0.093^{*}	-0.030	-0.229^{*}
	(-3.03)	(-2.04)	(-0.51)	(-2.07)
Urban	-0.083	0.052	0.028	0.080
	(-0.92)	(0.90)	(0.39)	(0.73)
Reference category: 1945				
1944	-0.088	0.084	0.009	-0.163
	(-1.12)	(1.65)	(0.16)	(-1.76)
1946	-0.214^{**}	0.049	-0.074	-0.214^{*}
	(-2.62)	(0.92)	(-1.21)	(-2.15)
1947	-0.312^{***}	-0.028	-0.086	-0.200
	(-3.50)	(-0.48)	(-1.28)	(-1.83)
Famine exposure				
PN	-0.075	-0.029	0.083	0.035
	(-1.05)	(-0.64)	(1.71)	(0.45)
T3	-0.055	0.155^{**}	-0.018	-0.034
	(-0.72)	(3.05)	(-0.33)	(-0.38)
T2	-0.080	0.002	0.114^{*}	-0.094
	(-1.06)	(0.04)	(2.15)	(-1.10)
T1	-0.076	0.065	0.001	-0.150
	(-0.84)	(1.04)	(0.02)	(-1.39)
PC	-0.105	-0.125^{**}	0.002	0.153^{*}
	(-1.47)	(-2.67)	(0.05)	(2.01)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: Simple probit COPD

COPD	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.522^{***}	-0.730^{***}	-0.846^{***}	-0.833^{***}
	(-6.26)	(-14.47)	(-15.19)	(-10.12)
BMI	-0.001	0.016^{**}	0.014^{*}	0.017
	(-0.10)	(2.82)	(2.28)	(1.75)
Height	-0.002	-0.003	-0.002	-0.008^{*}
	(-0.57)	(-1.52)	(-0.81)	(-2.55)
Lower than fit health	0.110^{*}	0.204^{***}	0.213^{***}	0.278^{***}
	(2.10)	(5.87)	(5.96)	(5.33)
Lower than fit psych	0.039	0.100^{**}	0.025	0.045
	(0.96)	(3.23)	(0.72)	(0.86)
Reference category: White collar				
Professional	0.097	0.087	0.054	-0.061
	(1.26)	(1.94)	(1.41)	(-1.38)
Farm owner	-0.053	0.070	-0.133	0.095
	(-0.42)	(1.15)	(-1.29)	(0.60)
Skilled	0.036	-0.013	0.037	-0.072
	(0.66)	(-0.42)	(1.06)	(-0.99)
Unskilled	0.072	0.074	0.077	-0.151
	(1.21)	(1.89)	(1.60)	(-1.30)
Unknown	0.013	0.128^{*}	0.075	-0.056
	(0.15)	(2.33)	(1.16)	(-0.51)
Reference category: Selected city				
Non-urban	-0.108	-0.121^{**}	-0.043	0.016
	(-1.57)	(-2.83)	(-0.78)	(0.18)
Urban	-0.093	-0.191^{***}	-0.209^{**}	-0.029
	(-1.09)	(-3.34)	(-2.93)	(-0.29)
Reference category: 1945				
1944	0.019	0.036	0.054	-0.101
	(0.26)	(0.75)	(0.99)	(-1.26)
1946	-0.027	-0.041	-0.049	-0.016
	(-0.34)	(-0.84)	(-0.85)	(-0.19)
1947	-0.070	-0.080	-0.025	-0.193^{*}
	(-0.84)	(-1.46)	(-0.39)	(-2.04)
Famine exposure				
PN	-0.050	-0.128^{**}	-0.053	0.028
	(-0.74)	(-2.98)	(-1.15)	(0.41)
T3	0.010	0.036	0.003	-0.021
	(0.13)	(0.75)	(0.06)	(-0.26)
T2	-0.078	0.079	0.017	-0.010
	(-1.07)	(1.71)	(0.35)	(-0.13)
T1	0.056	-0.067	-0.014	-0.091
	(0.65)	(-1.16)	(-0.22)	(-0.98)
PC	-0.058	-0.002	0.055	-0.079
	(-0.87)	(-0.04)	(1.21)	(-1.20)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Simple probit depression

Depression	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.523^{***}	-0.708^{***}	-0.657^{***}	-0.633^{***}
	(-6.25)	(-14.11)	(-12.24)	(-7.98)
BMI	0.001	-0.003	-0.001	0.000
	(0.10)	(-0.59)	(-0.24)	(0.04)
Height	0.001	-0.001	-0.004	0.001
	(0.33)	(-0.27)	(-1.67)	(0.18)
Lower than fit health	0.055	0.055	0.100^{**}	0.090
	(1.05)	(1.57)	(2.82)	(1.75)
Lower than fit psych	0.062	0.154^{***}	0.156^{***}	0.103^{*}
	(1.51)	(5.05)	(4.66)	(2.06)
Reference category: White collar				
Professional	0.060	0.003	0.019	0.060
	(0.77)	(0.07)	(0.52)	(1.43)
Farm owner	-0.137	-0.052	0.004	-0.027
	(-1.07)	(-0.85)	(0.04)	(-0.16)
Skilled	0.014	0.006	-0.021	-0.079
	(0.26)	(0.19)	(-0.63)	(-1.11)
Unskilled	-0.082	0.018	-0.031	0.101
	(-1.36)	(0.47)	(-0.66)	(0.96)
Unknown	0.038	0.069	-0.044	-0.098
	(0.45)	(1.25)	(-0.69)	(-0.91)
Reference category: Selected city				
Non-urban	-0.023	-0.088^{*}	0.008	-0.166
	(-0.34)	(-2.09)	(0.15)	(-1.95)
Urban	0.052	-0.069	-0.070	-0.053
	(0.61)	(-1.25)	(-1.05)	(-0.56)
Reference category: 1945				
1944	0.004	-0.028	-0.119^{*}	-0.164^{*}
	(0.05)	(-0.58)	(-2.26)	(-2.12)
1946	0.031	0.025	-0.109	-0.207^{*}
	(0.40)	(0.52)	(-1.95)	(-2.49)
1947	-0.020	-0.001	-0.133^{*}	-0.161
	(-0.24)	(-0.002)	(-2.18)	(-1.79)
Famine exposure				
PN	-0.145^{*}	-0.028	-0.048	-0.079
	(-2.08)	(-0.66)	(-1.07)	(-1.20)
T3	-0.028	0.035	-0.031	-0.116
	(-0.38)	(0.73)	(-0.62)	(-1.54)
T2	0.058	0.006	-0.062	-0.081
	(0.80)	(0.13)	(-1.27)	(-1.12)
T1	-0.010	-0.056	-0.150^{*}	-0.282^{**}
	(-0.12)	(-0.95)	(-2.43)	(-3.12)
PC	-0.060	-0.074	0.049	0.078
	(-0.90)	(-1.73)	(1.10)	(1.23)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8: Simple probit heartburn

Heartburn	Primary	Lower vocational	Lower secondary	Higher
Constant	0.259**	0.146**	0.013	0.057
	(3.21)	(3.18)	(0.27)	(0.79)
BMI	0.008	0.011*	0.008	0.020*
	(0.95)	(2.11)	(1.52)	(2.43)
Height	−0.004	−0.002	−0.002	−0.006*
	(−1.29)	(−1.13)	(−0.94)	(−2.29)
Lower than fit health	0.138**	0.080*	0.111***	0.047
	(2.65)	(2.41)	(3.35)	(1.00)
Lower than fit psych	−0.047	0.036	0.052	−0.032
	(−1.18)	(1.25)	(1.67)	(−0.71)
Reference category: White collar				
Professional	0.106	−0.024	−0.009	0.017
	(1.39)	(−0.58)	(−0.26)	(0.45)
Farm owner	−0.120	−0.064	−0.102	0.054
	(−1.00)	(−1.17)	(−1.18)	(0.38)
Skilled	0.003	−0.039	0.000	−0.061
	(0.06)	(−1.35)	(0.02)	(−0.97)
Unskilled	0.087	0.030	−0.001	0.043
	(1.50)	(0.83)	(−0.02)	(0.44)
Unknown	−0.020	0.062	0.079	−0.128
	(−0.24)	(1.21)	(1.36)	(−1.35)
Reference category: Selected city				
Non-urban	−0.049	−0.025	0.049	0.014
	(−0.74)	(−0.64)	(1.01)	(0.18)
Urban	0.017	−0.123*	−0.071	−0.145
	(0.21)	(−2.45)	(−1.18)	(−1.72)
Reference category: 1945				
1944	0.032	0.049	0.023	−0.123
	(0.44)	(1.14)	(0.48)	(−1.75)
1946	−0.011	0.022	−0.060	−0.211**
	(−0.15)	(0.49)	(−1.17)	(−2.82)
1947	−0.051	−0.030	−0.085	−0.176*
	(−0.64)	(−0.62)	(−1.52)	(−2.16)
Famine exposure				
PN	−0.073	0.013	−0.014	−0.074
	(−1.10)	(0.33)	(−0.35)	(−1.26)
T3	−0.017	−0.010	−0.017	−0.092
	(−0.23)	(−0.22)	(−0.37)	(−1.34)
T2	0.100	0.028	−0.006	−0.064
	(1.41)	(0.65)	(−0.13)	(−0.99)
T1	−0.036	−0.011	−0.035	−0.193*
	(−0.42)	(−0.21)	(−0.62)	(−2.38)
PC	0.013	−0.007	0.038	0.058
	(0.20)	(−0.19)	(0.94)	(1.02)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9: Simple probit hypertension

Hypertension	Primary	Lower vocational	Lower secondary	Higher
Constant	0.255**	0.242***	0.224***	0.025
	(3.17)	(5.24)	(4.52)	(0.35)
BMI	0.033***	0.043***	0.049***	0.056***
	(4.05)	(8.03)	(8.64)	(6.55)
Height	0.001	−0.005**	−0.006**	−0.009**
	(0.24)	(−2.68)	(−3.14)	(−3.28)
Lower than fit health	0.069	0.034	0.073*	0.045
	(1.34)	(1.02)	(2.19)	(0.94)
Lower than fit psych	−0.025	0.046	0.016	0.054
	(−0.63)	(1.58)	(0.52)	(1.18)
Reference category: White collar				
Professional	0.050	−0.071	−0.002	−0.065
	(0.66)	(−1.72)	(−0.04)	(−1.69)
Farm owner	−0.115	−0.151**	−0.159	−0.044
	(−0.96)	(−2.73)	(−1.82)	(−0.31)
Skilled	0.023	−0.030	0.041	0.061
	(0.45)	(−1.05)	(1.32)	(0.98)
Unskilled	0.030	−0.026	0.002	0.060
	(0.52)	(−0.72)	(0.04)	(0.61)
Unknown	−0.008	0.032	0.153**	0.078
	(−0.09)	(0.61)	(2.61)	(0.83)
Reference category: Selected city				
Non-urban	0.030	−0.081*	−0.060	0.028
	(0.46)	(−2.09)	(−1.23)	(0.37)
Urban	−0.053	−0.111*	0.048	−0.116
	(−0.65)	(−2.19)	(0.80)	(−1.37)
Reference category: 1945				
1944	0.0125	0.035	−0.053	0.042
	(0.17)	(0.80)	(−1.09)	(0.60)
1946	−0.057	−0.089*	−0.225***	−0.054
	(−0.77)	(−1.98)	(−4.38)	(−0.73)
1947	−0.137	−0.096	−0.160**	−0.196*
	(−1.71)	(−1.94)	(−2.86)	(−2.40)
Famine exposure				
PN	−0.004	−0.037	0.002	−0.042
	(−0.06)	(−0.95)	(0.05)	(−0.71)
T3	0.129	0.096*	−0.027	0.040
	(1.78)	(2.17)	(−0.58)	(0.58)
T2	0.064	−0.016	−0.002	0.021
	(0.91)	(−0.36)	(−0.05)	(0.32)
T1	−0.096	−0.062	−0.124*	−0.040
	(−1.14)	(−1.15)	(−2.21)	(−0.50)
PC	−0.051	−0.046	−0.035	−0.021
	(−0.79)	(−1.15)	(−0.87)	(−0.37)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: Simple probit cardiac diseases

Cardiac diseases	Primary	Lower vocational	Lower secondary	Higher
Constant	-0.493^{***}	-0.638^{***}	-0.695^{***}	-0.943^{***}
	(-5.89)	(-12.76)	(-12.79)	(-11.28)
BMI	0.039^{***}	0.038^{***}	0.042^{***}	0.038^{***}
	(4.78)	(6.81)	(7.02)	(3.96)
Height	0.001	-0.001	-0.002	0.004
	(0.18)	(-0.51)	(-0.87)	(1.20)
Lower than fit health	0.025	0.062	0.106^{**}	0.034
	(0.46)	(1.75)	(2.91)	(0.61)
Lower than fit psych	0.053	0.067^{*}	0.039	0.031
	(1.29)	(2.16)	(1.11)	(0.58)
Reference category: White collar				
Professional	-0.137	0.031	0.011	0.023
	(-1.73)	(0.70)	(0.28)	(0.51)
Farm owner	-0.199	-0.112	-0.223^{*}	0.306^{*}
	(-1.53)	(-1.80)	(-2.14)	(2.01)
Skilled	-0.040	-0.013	0.016	-0.066
	(-0.74)	(-0.42)	(0.46)	(-0.88)
Unskilled	-0.073	0.092^{*}	0.051	0.051
	(-1.21)	(2.38)	(1.08)	(0.46)
Unknown	-0.040	0.104	0.144^{*}	0.067
	(-0.48)	(1.90)	(2.31)	(0.61)
Reference category: Selected city				
Non-urban	-0.021	-0.092^{*}	-0.057	0.111
	(-0.29)	(-2.19)	(-1.04)	(1.30)
Urban	0.153	-0.082	-0.006	-0.034
	(1.81)	(-1.48)	(-0.09)	(-0.33)
Reference category: 1945				
1944	0.011	-0.001	-0.043	0.063
	(0.14)	(-0.20)	(-0.81)	(0.79)
1946	-0.115	-0.061	-0.162^{**}	-0.122
	(-1.47)	(-1.24)	(-2.86)	(-1.40)
1947	-0.210^{*}	-0.159^{**}	-0.189^{**}	-0.131
	(-2.48)	(-2.92)	(-3.03)	(-1.36)
Famine exposure				
PN	-0.018	-0.135^{**}	-0.043	0.030
	(-0.27)	(-3.17)	(-0.95)	(0.44)
T3	0.105	0.021	-0.069	0.049
	(1.42)	(0.44)	(-1.35)	(0.62)
T2	0.002	-0.032	0.012	-0.027
	(0.03)	(-0.70)	(0.24)	(-0.36)
T1	-0.050	-0.019	-0.084	0.028
	(-0.57)	(-0.33)	(-1.38)	(0.29)
PC	0.097	-0.080	0.038	0.074
	(1.43)	(-1.85)	(0.85)	(1.09)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11: Simple probit ischemic heart disease

Ischemic heart disease	Primary	Lower vocational	Lower secondary	Higher
Constant	-1.007^{***}	-1.032^{***}	-1.129^{***}	-1.400^{***}
	(-10.82)	(-18.23)	(-18.07)	(-13.76)
BMI	0.015	0.027^{***}	0.019^{**}	0.035^{**}
	(1.68)	(4.22)	(2.78)	(3.02)
Height	-0.010^{**}	-0.013^{***}	-0.014^{***}	-0.006
	(-2.87)	(-5.81)	(-5.78)	(-1.48)
Lower than fit health	-0.083	0.001	0.037	-0.030
	(-1.38)	(0.03)	(0.87)	(-0.45)
Lower than fit psych	0.028	0.043	0.061	0.110
	(0.61)	(1.22)	(1.56)	(1.76)
Reference category: White collar				
Professional	-0.042	0.037	0.006	0.089
	(-0.47)	(0.72)	(0.13)	(1.67)
Farm owner	-0.054	-0.152^{*}	-0.203	0.137
	(-0.38)	(-2.07)	(-1.64)	(0.74)
Skilled	0.082	0.009	0.008	-0.076
	(1.37)	(0.24)	(0.19)	(-0.83)
Unskilled	0.028	0.102^{*}	0.096	-0.235
	(0.43)	(2.33)	(1.79)	(-1.50)
Unknown	-0.010	0.135^{*}	0.169^{*}	0.187
	(-0.10)	(2.22)	(2.43)	(1.51)
Reference category: Selected city				
Non-urban	0.000	-0.109^{*}	-0.028	0.068
	(0.01)	(-2.24)	(-0.44)	(0.67)
Urban	0.131	-0.042	0.028	-0.135
	(1.41)	(-0.67)	(0.36)	(-1.08)
Reference category: 1945				
1944	0.083	-0.021	-0.028	0.102
	(1.00)	(-0.39)	(-0.45)	(1.02)
1946	-0.049	-0.050	-0.122	0.006
	(-0.57)	(-0.90)	(-1.87)	(0.05)
1947	-0.081	-0.138^{*}	-0.118	0.006
	(-0.86)	(-2.22)	(-1.65)	(0.05)
Famine exposure				
PN	0.023	-0.108^{*}	-0.030	-0.059
	(0.31)	(-2.24)	(-0.58)	(-0.73)
T3	0.114	0.028	0.027	0.140
	(1.41)	(0.52)	(0.47)	(1.44)
T2	0.091	0.012	0.010	-0.040
	(1.13)	(0.23)	(0.17)	(-0.44)
T1	-0.020	-0.048	-0.044	0.106
	(-0.20)	(-0.74)	(-0.62)	(0.93)
PC	0.106	-0.003	0.094	-0.097
	(1.41)	(-0.07)	(1.82)	(-1.19)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix B

Table B1 reports the average of each observed covariates included in vector C, for the individuals linked to medication use, allowing to reproduce the predicted probabilities at means over a significant interval of the latent intelligence (Figures 1-8). Recall that, in regressions, both height and BMI are mean centered.

Table B1: Mean of covariates in the medication use probit.

Variable	Mean	Variable	Mean
BMI	0.000	Non-urban	0.115
Height	0.000	Urban	0.055
Lower than fit health	0.156	1944	0.175
Lower than fit psych	0.214	1946	0.297
Professional	0.165	1947	0.126
Farm owner	0.035	PN	0.184
Skilled	0.271	T3	0.183
Unskilled	0.125	T2	0.152
Unknown	0.054	T1	0.103
		PC	0.173

Bibliography

[1] Bagger J., Birchenall J.A., Mansour H., Urzúa S. (2013), Education, Birth Order, and Family Size, *IZA*, DP No. 7454.

[2] Bijwaard G.E., van Kippersluis H., Veenman J. (2015), Education and health: the role of cognitive ability, *Journal of Health Economics*, 42: 29–43.

[3] Bijwaard G.E., van Poppel F., Ekamper P., Lumey L.H. (2015), Gains in life expectancy associated with higher education in men, *PLoS ONE*, 10(10): e0141200.

[4] Clark D., Royer H. (2013), The Effect of Education on Adult Mortality and Health: Evidence from Britain, *The American Economic Review*, 103(6): 2087-2120.

[5] Conti G., Heckman J.J., Urzúa S. (2010), The Education-Health Gradient, *The American Economic Review*, 100(2): 234-38.

[6] Doornbos G., Kromhout D. (1990), Educational level and mortality in a 32-year follow-up study of 18-year-old men in the Netherlands, *International Journal of Epidemiology*, 19: 374–379.

[7] Ekamper P., van Poppel F., Stein A.D., Lumey L.H. (2014), Independent and additive association of prenatal famine exposure and intermediary life conditions with adult mortality between age 18-63 years, *Social Science & Medicine*, 119: 232-239.

[8] Halfon P., Eggli Y., Decollogny A., et al. (2013), Disease identification based on ambulatory drugs dispensation and in-hospital ICD-10 diagnoses: a comparison, *BMC Health Services Research*, 13: 453.

[9] Heckman J.J., Humphries J.E., and Veramendi G. (2016), Returns to education: The causal effects of education on earnings, health and smoking, *Journal of Political Economy*, (forthcoming).

[10] Heckman J.J., Tobias J.L., Vytlacil E. (2000), Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Returns to Schooling, *National Bureau of Economic Research*, Working Paper 7950.

[11] Jackson M. (2009), Understanding Links between Adolescent Health and Educational Attainment, *Demography*, 46(4): 671–94.

[12] Lochner L. (2011), Non-production benefits of education: Crime, health, and good citizenship. *National Bureau of Economic Research*, Working Paper 16722.

[13] Long J.S. (2009), Group comparisons in logit and probit using predicted probabilities, Working Paper.

[14] Lumey L.H., Stein A.D., Kahn H.S. (2009), Food restriction during gestation and impaired fasting glucose or glucose tolerance and type 2 diabetes mellitus in adulthood: evidence from the Dutch hunger winter families study, *Journal of Developmental Origins of Health and Disease*, 1(S1), S164.

[15] Montez J.K., Hayward M.D., Brown D.C., Hummer D.A. (2009), Why Is the Educational Gradient of Mortality Steeper for Men? *The Journals of Gerontology*, 64B (5): 625-634.

[16] van Ooijen R., Alessie R., Knoef M. (2015), Health status over the life cycle, *Netspar*, DP 10/2015-062.

[17] Shipley B. (2000), *Cause and Correlation in Biology*, Cambridge University Press, Cambridge, U.K.

[18] Smith J.P. (2007), Nature and causes of trends in male diabetes prevalence, undiagnosed diabetes, and the socioeconomic status health gradient, *PNAS,* 104 (33).