

Brunello, Giorgio; Fort, Margherita; Schneeweis, Nicole; Winter-Ebmer, Rudolf

**Working Paper**

## The causal effect of education on health: What is the role of health behaviors?

IZA Discussion Papers, No. 5944

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Brunello, Giorgio; Fort, Margherita; Schneeweis, Nicole; Winter-Ebmer, Rudolf (2011) : The causal effect of education on health: What is the role of health behaviors?, IZA Discussion Papers, No. 5944, Institute for the Study of Labor (IZA), Bonn, <https://nbn-resolving.de/urn:nbn:de:101:1-201109288874>

This Version is available at:

<https://hdl.handle.net/10419/55058>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 5944

**The Causal Effect of Education on Health:  
What is the Role of Health Behaviors?**

Giorgio Brunello  
Margherita Fort  
Nicole Schneeweis  
Rudolf Winter-Ebmer

August 2011

# The Causal Effect of Education on Health: What is the Role of Health Behaviors?

**Giorgio Brunello**

*University of Padua, CESifo and IZA*

**Margherita Fort**

*University of Bologna and CHILD*

**Nicole Schneeweis**

*University of Linz*

**Rudolf Winter-Ebmer**

*University of Linz, CEPR, IHS and IZA*

Discussion Paper No. 5944

August 2011

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The Causal Effect of Education on Health: What is the Role of Health Behaviors?\***

In this paper we investigate the contribution of health related behaviors to the education gradient, using an empirical approach that addresses the endogeneity of both education and behaviors in the health production function. We apply this approach to a multi-country data set, which includes 12 European countries and has information on education, health and health behaviors for a sample of individuals aged 50+. Focusing on self reported poor health as our health outcome, we find that education has a protective role both for males and females. When evaluated at the sample mean of the dependent variable, one additional year of education reduces self-reported poor health by 7.1% for females and by 3.1% for males. Health behaviors – measured by smoking, drinking, exercising and the body mass index – contribute to explaining the gradient. We find that the effects of education on smoking, drinking, exercising and eating a proper diet account for at most 23% to 45% of the entire effect of education on health, depending on gender.

JEL Classification: J1, I12, I21

Keywords: health, education, health behaviors, Europe

Corresponding author:

Giorgio Brunello  
Department of Economics  
University of Padova  
Via del Santo 33  
35100 Padova  
Italy  
E-mail: [giorgio.brunello@unipd.it](mailto:giorgio.brunello@unipd.it)

---

\* We would like to thank the participants to seminars in Bologna, Bressanone, Catanzaro, Firenze, Hangzhou, Linz, Nurnberg, Padova, Regensburg and Wurzburg for comments and suggestions on an earlier version of the paper. We acknowledge the financial support of Fondazione Cariparo, MIUR-FIRB 2008 project RBF089QQC-003-J31J10000060001 and the Austrian Science Funds ("The Austrian Center for Labor Economics and the Analysis of the Welfare State"). The SHARE data collection has been primarily funded by the European Commission through the 5th, 6th and 7th framework programme, as well as from the U.S. National Institute on Aging and other national Funds. The usual disclaimer applies.

# 1 Introduction

The relationship between education and health - the "education gradient" - is widely studied. There is abundant evidence that a gradient exists (Cutler and Lleras-Muney, 2010). Yet less is known as to why education might be related to health. In this paper we explore the contribution of health related behaviors (shortly, behaviors) - which we measure with smoking, drinking, exercising and having a poor diet - to the education gradient. To do so, we decompose the gradient into two parts: a) the part mediated by health behaviors; b) a residual, which includes for instance stress reduction, better decision making, better information collection, healthier employment and better neighborhoods (Lochner, 2011)<sup>1</sup>.

We are not the first to investigate the mediating role of health behaviors. As recently pointed out by Lochner (2011), a problem with the existing empirical literature is that most contributions fail to address the endogeneity of education and behaviors in health regressions: there are possibly many confounding factors which influence both education and behaviors, on the one hand, and health outcomes, on the other hand. While some studies have dealt with endogenous education, our approach is novel because we address the endogeneity of both education and behaviors in the health production function, and therefore can give a causal interpretation to our estimates.

Our identification strategy - based on the work by Card and Rothstein (2007) - allows us to estimate average education effects for an individual randomly picked from the population. Using a cross-country dataset, where we have a rich set of parental and early life information, this strategy combines selection on observables and fixed effects assumptions to estimate the parameters of both a dynamic health equation, which depends on education and lagged health behaviors, and a static health equation, where health depends only on education. The effect of education on health in the second equation is the education gradient (shortly, the gradient), i.e. the total effect of education on health that results from both mediated and residual effects of education.

We compare the estimates of the gradient obtained following the strategy outlined above with those obtained with a completely different methodology, instrumental variables (IV) estimation, where the key exogenous variation is provided by the changes in compulsory school leaving ages across countries and birth cohorts. While the IV strategy generates causal estimates that are internally valid for individuals affected by mandatory schooling laws (*compliers*), it cannot be used for the decomposition of the education gradient, because of the lack of valid and relevant instruments for behaviors.

We apply this approach to a multi-country data set, which includes 12 European countries (Austria, Belgium, Denmark, England, France, Germany, Greece, Italy, the

---

<sup>1</sup>The residual also includes the contribution of unmeasured behaviors.

Netherlands, Spain, Sweden and Switzerland) and has information on education, health and health behaviors for a sample of males and females aged 50+. By focusing on older individuals, we consider the long term effects of education on health. These data are drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE) and from the English Longitudinal Study of Ageing (ELSA). Both surveys are modeled following the US Health and Retirement Study.

Focusing on self-reported poor health as our health outcome, we find that education has a protective role both for males and females, although effects for females are typically somewhat higher. When evaluated at the sample mean of the dependent variable, one additional year of education reduces self-reported poor health by 7.1% for females and by 3.1% for males. These effects are smaller than those found by others. Our explanation is that we use a sample of older individuals (50+) typically done in the literature, and that the protective role of education on health declines with age.

Our qualitative findings are robust to the choice of the identification strategy. The absolute size of the gradient, however, is largest when we focus on the compliers to compulsory school reforms. For this sub-group we find that, when evaluated at the sample mean of the dependent variable, one additional year of education reduces self perceived poor health by 16.5% and 12.1% for males and females respectively. Since compliers are typically drawn among those with lower education, our findings suggest that improving the education of this group is particularly rewarding in terms of better self perceived health.

There is also evidence that health behaviors - measured by smoking, drinking, exercising and the body mass index - contribute to explaining the gradient. The size of this contribution is larger when we consider the entire history of behaviors rather than only behaviors in the immediate past. In the former case, we find that the effects of education on smoking, drinking, exercising and eating a proper diet account for at most 23% to 45% of the entire effect of education on health, depending on gender. The largest part of the gradient, however, remains unaccounted for. Potential candidates include direct effects of education on health as well as indirect effects operating through unobserved health behaviors, wealth and cognitive abilities.

The paper is organized as follows: Section 2 is a brief review of the relevant literature. The theoretical model is presented in Section 3, and our empirical strategy is discussed in Section 4. Section 5 describes the data. The empirical results are discussed in Section 6. Conclusions follow.

## 2 Review of the Literature

As recently reviewed by Lochner (2011), the empirical research on the causal effect of education on health has produced so far mixed results. This literature typically focuses on the impact on self-reported health and on single countries (Clark and Royer (2010), Juerges et al. (2009), Silles (2009), Adams (2002), Arendt (2005), Arendt (2008), Albouy and Lequien (2009)) and identifies the effect of education on health by using the exogenous variation generated by changes in mandatory schooling laws. Some of these studies find that education improves self reported health (Mazumder (2008) for the US and Silles (2009) for the UK). Others find no effect (Clark and Royer (2010), Oreopolous (2007), Braakmann (2011) and Juerges et al. (2009) for the UK, Arendt (2005) for Denmark). While Silles (2009) finds that education reduces self reported long term illness in the UK, Kempter et al. (2011), find a protective role of education for German males but not for German females<sup>2</sup>.

There are many possible channels through which education may improve health. Lochner (2011) lists the following: stress reduction, better decision making and/or better information gathering, higher likelihood of having health insurance, healthier employment, better neighborhoods and peers and healthier behaviors.<sup>3</sup> The contribution of behaviors, which include smoking, drinking and eating calorie-intensive food, has been examined in the economic and sociological literature, starting with the contribution by Ross and Wu (1995)<sup>4</sup>. These authors use US data, regress measures of health on income, social resources and behaviors and treat both behaviors and education as exogenous. They find that behaviors explain less than 10% of the education gradient.

Cutler et al. (2008) discuss possible mechanisms underlying the education gradient. Using data from the NHIS survey in the US, they find that behaviors account for over 40% of the effect of education on mortality in their sample of non-elderly Americans. A problem with these studies is that they fail to consider the endogeneity of both education and behaviors in a health equation which includes both. In the study closest to the current paper, Contoyannis and Jones (2004) partly address this concern by explicitly modeling the optimal choice of health behaviors. They jointly estimate a health equation - where health depends on education and behaviors - and separate behavior equations - where behaviors depend on education - by FIML (Full Information Maximum Likelihood), treating education as exogenous. Using Canadian data,

---

<sup>2</sup>While most studies consider self reported health, Powdthavee (2010), examines the effects of education on hypertension, as determined from blood pressure measurements, Meghir et al. (2011) study mortality in Sweden and Brunello et al. (2011) study the effects on several chronic diseases.

<sup>3</sup>Conti et al. (2010) argue that non-cognitive skills may be an important factor as well.

<sup>4</sup>See the reviews by Feinstein et al. (2006) and Cawley and Ruhm (2011).

they show that the contribution of lagged (7 years earlier) behaviors to the education gradient varies between 23% to 73%, depending on whether behaviors are treated as exogenous or endogenous.

We summarize the existing evidence as follows: first, the available empirical evidence on the causal effect of education on health is mixed at best and covers a rather limited set of countries (US, UK, Canada, Germany, Denmark and France); second, the estimated contribution of behaviors to the education gradient varies substantially across the few available studies, depending on model specification and identification strategy.<sup>5</sup>

We contribute to this literature in several directions. Our study is the first to cover a substantial number of European countries (12), using a multi-country dataset which includes also Southern European countries, which have not been studied before. We are also the first to offer an identification strategy which addresses the endogeneity of both education and health behaviors in the health production function. The estimates of the education gradient based on this strategy are compared with those obtained with a more conventional IV strategy, which uses the exogenous variation across countries and cohorts induced by changes in mandatory school leaving age. Finally, we distinguish explicitly between the short run and long run mediating effects of health behaviors. While the former only include the effects of current or lagged behaviors, the latter takes into account the contribution of the entire history of behaviors. This qualification is empirically relevant as we show in section 6.

### 3 The Model

Following Grossman (1972), Rosenzweig and Schultz (1983) and Contoyannis and Jones (2004), assume that individuals have preference orderings over their own poor health  $H$  and two bundles of goods,  $C$  and  $B$ , where only the latter affects health. The vector  $B$  includes risky health behaviors or habits - such as smoking, the use of alcohol or drugs, unprotected sex, excessive calorie intake and poor exercise - which increase the utility from consumption but damage health<sup>6</sup>. Utility  $U(C, B, H)$  is concave in its arguments and the marginal utility of consumption ( $U_C$  and  $U_B$ ) varies with health<sup>7</sup>. Reflecting the view that better educated individuals have access to higher income and

---

<sup>5</sup>See also Stowasser et al. (2011) for a discussion on causality issues between socio-economic status in general and health.

<sup>6</sup>See the discussion in Feinstein et al. (2006)

<sup>7</sup>The sub-scripts are for partial derivatives. The relationship between health and the marginal utility of consumption is not clear ex-ante. On the one hand, the latter may decline with deteriorating health, because several consumption goods are complements to good health. On the other hand, deteriorating health may increase the marginal utility of consumption. ". . . as other consumption goods - such as prepared meals or assistance with self-care - are substitutes for health . . ." (Finkelstein et al., 2008)



can therefore extract higher utility from better health and a longer life, we assume that the marginal utility of (poor) health declines when individual education  $E$  increases, that is  $U_{HE} < 0$ <sup>8</sup>.

The stock of individual poor health  $H$  is positively affected by behaviors  $B$  and negatively affected by individual education  $E$ . As reviewed by Lochner (2011), channels through which education may improve health include stress reduction, better decision making, healthier and safer employment, healthier neighborhoods and peers. Poor health  $H$  depends also on a vector of unobservables  $\mu$ , which include both parental and job characteristics (see Park, 2008). Using a linear specification, the health production function is given by

$$H = \alpha B - \beta E + \gamma \mu \quad (1)$$

Rational individuals maximize their utility with respect to consumption, subject to the health production function and to the budget constraint, defined as follows<sup>9</sup>

$$pC + B = Y(E, X) \quad (2)$$

where  $Y$  is income, which varies with education and a vector of observable controls  $X$ ,  $p$  is the vector of consumption prices for goods  $C$  and the prices of  $B$  are normalized to 1. Assuming that an internal solution exists, the necessary conditions for a maximum are

$$U_C - \lambda p = 0 \quad (3)$$

$$U_B + \alpha U_H - \lambda = 0 \quad (4)$$

where  $\lambda$  is the Lagrange multiplier. Concavity of the utility function implies  $U_{HH} < 0$ . Moreover, Finkelstein et al. (2008), find that the marginal utility of consumption declines when health deteriorates. Therefore,  $U_{CH} < 0$  and  $U_{BH} < 0$ . By totally differentiating (3) and (4) and using (1) we obtain that higher education reduces health damaging behaviors if the following condition holds<sup>10</sup>

$$|U_{HE}| > \beta \left[ \left( \frac{|U_{BH}|}{\alpha} - \frac{|U_{CH}|}{p\alpha} \right) + |U_{HH}| \right] \quad (5)$$

---

<sup>8</sup>As argued by Cutler and Lleras-Muney (2006), the higher weight placed on health by the better educated could reflect the higher value of the future: "...if education provides individuals with a better future along several dimensions - people may be more likely to invest in protecting that future". (p.15)

<sup>9</sup>Rosenzweig and Schultz (1983), and Contoyannis and Jones (2004), use a similar formulation.

<sup>10</sup>We assume that the second order conditions for a maximum hold. Condition (5) also ensures that higher education increases consumption  $C$ . When utility is separable in consumption and health - as in Cutler et al. (2003)

$$U(C, B, H) = U(C) + \Omega(B) - h(E)H$$

condition (5) is verified if  $h_E(E) > 0$ .

The optimal consumption plan in implicit form is given by

$$C = C(E, p, \mu, X) \quad (6)$$

$$B = B(E, p, \mu, X) \quad (7)$$

Using (7) in (1) and in the utility function yields the "reduced form" health equation

$$H = H(E, p, \mu, X) \quad (8)$$

and the indirect utility function  $V = V(E, p, \mu, X)$ . The marginal effect of education on health in (8) is the "education gradient" ( $HEG$ ). Assuming that the cost of education  $\Gamma(E, Z)$ , where  $Z$  is a vector of cost of education shifters, is convex in the years of education, optimal education is given by

$$V_E(E, p, \mu, X) = \Gamma_E(E, Z) \quad (9)$$

### 3.1 The Contribution of Health Behaviors to the Education Gradient: Current or Lagged Behaviors

In the empirical literature (Ross and Wu (1995) or Cutler et al. (2008)) the contribution of health behaviors to the education gradient is evaluated by using either current or lagged behaviors in equation (1). The lag is often justified with the view that the impact of health behaviors on health requires time. In this case, and omitting unobservables  $\mu$  for the sake of simplicity, the health production function (1) can be re-written as

$$H_t = \alpha B_{t-1} - \beta E \quad (10)$$

where  $t$  is time, and the education gradient can be decomposed into: a) the effect operating via health behaviors lagged once  $B_{t-1}$ ; b) a residual effect. The ratio between a) and the overall effect measures the relative contribution of health behaviors lagged once to the education gradient.

To illustrate with an example, assume that utility is given by  $U(C_t, B_t, H_t) = \Phi(C_t) + \Gamma(B_t) - h(E)H_t$  and let  $\rho$  be the discount factor. Under these assumptions, optimal behavior is  $B_t = B(E, X_t, p_t, \rho)$ . Ignoring for the time being the price vector  $p$ , the discount factor and the vector  $X$ , a linear approximation of this behavior is

$$B_t = \lambda_0 - \lambda_1 E \quad (11)$$

Substituting (11) into (10) yields

$$H_t = \alpha \lambda_0 - (\alpha \lambda_1 + \beta_1) E \quad (12)$$

The gradient is  $-(\alpha\lambda_1 + \beta_1)$  and the relative contribution of behaviors lagged once to the gradient is  $\frac{\alpha\lambda_1}{(\alpha\lambda_1 + \beta_1)}$ .

### 3.2 The Contribution of Health Behaviors to the Education Gradient: The History of Behaviors

By focusing on current or lagged behaviors, specification (10) explicitly assumes that previous lags do not contribute to current health conditional on behaviors observed in the previous period. To illustrate again with an example the implications of this assumption, let the "true" health production function be given by

$$H_t = k_0 + k_1 B_{t-1} + k_2 B_{t-2} + \dots + k_T B_{t-T} - \theta E \quad (13)$$

This function is more general than (10) because current health depends both on behaviors lagged once and on previous lags from  $t - 2$  to the initial period  $T$ . Ignoring again the price vector  $p$ , the discount factor and the vector  $X$ , a linear approximation of optimal behaviors is given by  $B_t = \sigma_0 - \sigma_1 E$ , combined with (13) yields

$$H_t = k_0 + k_1 B_{t-1} - [\sigma_1(k_2 + \dots + k_T) + \theta] E \quad (14)$$

When the health production function depends on risky health behaviors lagged 1 to  $T$ , the contribution of behaviors lagged once to the education gradient is  $\frac{\sigma_1 k_1}{[\sigma_1(k_1 + k_2 + \dots + k_T) + \theta]}$ , where the denominator includes both the effect of education on health conditional on behaviors  $\theta$  and the mediating effects of behaviors from lag 1 to  $T$ . This contribution differs from the contribution of health behaviors lagged 1 to  $T$ , which is given by  $\frac{\sigma_1(k_1 + k_2 + \dots + k_T)}{[\sigma_1(k_1 + k_2 + \dots + k_T) + \theta]}$ . If the parameters  $k_i$  are positive, ignoring the contribution of higher lags leads to under-estimating the overall mediating effect of risky health behaviors.

When the available data do not include information on behaviors from lag  $t - 2$  to lag  $T$ , as it happens in our case, an alternative approach is to adopt the dynamic health equation (see for instance Park and Kang (2008))

$$H_t = \pi B_{t-1} - \nu E + \phi H_{t-1} \quad (15)$$

which requires data only for periods  $t$  and  $t - 1$ . Under the assumptions that  $H_{t-T} = 0$  and  $\phi < 1$ , and ignoring again prices, the vector  $X$  and the discount factor,

equation (15) is equivalent to equation (13) when the following restrictions on the parameters hold

$$\begin{aligned}
k_1 &= \pi \\
\theta &= \nu \frac{1 - \phi^T}{1 - \phi} \\
k_2 &= \pi \phi \\
k_T &= \pi \phi^{T-1}
\end{aligned} \tag{16}$$

Since (15) can be written as (13)<sup>11</sup>, the linear approximation of optimal health behaviors is unchanged. Using this approximation into (13), taking into account constraints (16) and assuming that  $T$  is large ( $\phi^T \rightarrow 0$ ) we obtain

$$H_t = \pi B_{t-1} - \left[ \frac{\nu + \phi \sigma_1 \pi}{1 - \phi} \right] E \tag{17}$$

The education gradient - which includes also the mediating effect of health behaviors lagged once - is equal to  $-\frac{(\pi \sigma_1 + \nu)}{1 - \phi}$ . The relative contribution of health behaviors lagged once to the education gradient (short-run mediating effect, SRME) is

$$SRME = \frac{(1 - \phi)\pi \sigma_1}{(\pi \sigma_1 + \nu)} \tag{18}$$

The overall relative contribution of health behaviors (long-run mediating effect, LRME) to the education gradient adds to the contribution of health behaviors lagged once the contribution of lags from  $t - 2$  to  $T$ , and is equal to

$$LRME = \frac{\pi \sigma_1}{(\pi \sigma_1 + \nu)} \tag{19}$$

This implies that  $SRME = (1 - \phi)LRME$ . Under our assumptions,  $SRME$  underestimates  $LRME$ , and the degree of under-estimation is larger the higher is  $\phi$  (persistence of health status over time). Therefore, if we only estimate  $SRME$ , we may find a small contribution of health behaviors to the overall education gradient not because health behaviors have a small mediating effect but because we have ignored the contributions of health behaviors lagged more than once<sup>12</sup>.

---

<sup>11</sup>This is the case when the unobservable component  $\mu$  is either time invariant or follows an autoregressive process.

<sup>12</sup>If the education gradient is negative, sufficient conditions for the indicator  $LRME$  ( $SRME$ ) to fall within the range  $[0, 1]$  are  $\pi \sigma_1 \geq 0$  and  $\nu \geq 0$  ( $\phi \pi \sigma_1 + \nu \geq 0$ ). If the gradient is positive, these conditions also change sign.

### 3.3 Estimating the Short-Run and Long-Run Mediating Effects

One of the aims of this paper is to provide estimates of  $SRME$  and  $LRME$ . Our empirical strategy consists of estimating the parameters of both the dynamic health equation (15) and the "reduced form" health equation

$$H_t = \chi_o - \chi_1 E \quad (20)$$

where  $\chi_1 = \frac{(\pi\sigma_1 + \nu)}{1 - \phi}$ . Using these estimates, we can compute both

$$\widehat{\pi\sigma_1} = \widehat{\chi_1}(1 - \widehat{\phi}) - \widehat{\nu} \quad (21)$$

and

$$\widehat{LRME} = \frac{\widehat{\chi_1}(1 - \widehat{\phi}) - \widehat{\nu}}{\widehat{\chi_1}(1 - \widehat{\phi})} \quad (22)$$

$$\widehat{SRME} = (1 - \widehat{\phi})\widehat{LRME} \quad (23)$$

This strategy has the advantage that it only requires the estimation of two equations and the drawback that we cannot separately identify the mediating effect of each single health behavior. For that, we would need to estimate also Eq. (7) for each available behavior. We leave this development to future research.

## 4 The Empirical Strategy

The dynamic health equation (1) includes education  $E$ , behaviors  $B$  and unobservables  $\mu$ . On the one hand, optimal behaviors depend on  $\mu$  (see Eq. (7)). On the other hand, optimal education equalizes the marginal costs and the marginal benefits of education, and these benefits depend on  $\mu$ . Therefore, individual choice implies that both education and behaviors in Eq. (15) and education in the "reduced form" equation (20) are correlated with unobservables that affect health outcomes. Because of this, ordinary least squares estimates of either equation fail to uncover causal relationships. As remarked above, an important caveat for the empirical studies investigating the mediating effect of health behaviors on the education gradient is that they fail to consider these endogeneity problems (Lochner, 2011). In this paper, we address these problems in an attempt to give a causal interpretation both to the health-education gradient and to the mediating role of behaviors.

In the past few years, several papers have estimated the causal effect of education on health using the exogenous variation in educational attainment generated by changes in compulsory schooling. This instrumental variables (IV) approach can be used to

estimate the "reduced" form health equation (20). In principle, the same approach can also be applied to estimate the dynamic health production function (15), provided that we can find additional credible sources of exogenous variation which affect risky health behaviors without influencing individual health (conditional on behaviors). This is a very difficult task with the data at hand.<sup>13</sup> Therefore, we turn to the identification strategy suggested by Card and Rothstein (2007), which combines aggregation, selection on observables and fixed effects assumptions, to estimate both the dynamic health production function and the "reduced form" health equation. For the latter equation, we compare the results obtained following the Card and Rothstein (2007) approach to those obtained with a more standard IV approach, using changes in compulsory education as the relevant instrument. In the rest of this section, we illustrate the two approaches in turn.

#### 4.1 The Card-Rothstein approach

Consider the following empirical version of the dynamic health production function

$$H_{icgbt} = \alpha_0 + \alpha_1 B_{icgb(t-1)} + \alpha_2 E_{icg} + \alpha_3 X_{icgb} + \alpha_4 H_{icgbt-1} + \varepsilon_{icgbt} \quad (24)$$

where  $X$  is a vector of controls,  $B$  the vector of behaviors,  $\varepsilon$  is the error term,  $i$  denotes the individual,  $c$  the country,  $t$  the year of the interview,  $g$  gender, and  $b$  is the birth cohort.

Following Card and Rothstein (2007), we can decompose the error term in equation (24) as follows

$$\varepsilon_{icgbt} = u_{cgbt} + e_{icgbt} \quad (25)$$

where  $u_{cgbt}$  represents a common error component for individuals of the same gender  $g$  and birth cohort  $b$  in country  $c$  at time  $t$ ,  $e_{icgbt}$  is an individual specific idiosyncratic error component for which we assume

$$E[e_{icgbt} | b, g, c, t] = 0 \quad (26)$$

i.e., the individual specific error term has mean zero across individuals of the same gender, year of birth, country and time period.

---

<sup>13</sup>Using instruments like the price of alcohol or cigarettes has two main drawbacks. First, it would exploit only cross-sectional variation across different countries: indeed, all such potential instruments would influence all cohorts in one country alike. Second, it would prevent from the possibility to control for country fixed effects.

We aggregate individual data in cells identified by country, time, birth cohort and gender, define  $G$  as a gender dummy equal to 1 for females and to 0 for males and re-write Eq. (24) as follows

$$\begin{aligned}
H_{cbt} = & \alpha_0 + \alpha_0^F G + \alpha_1 B_{cb(t-1)} + \alpha_1^F G B_{cb(t-1)} + \alpha_2 E_{cb} + \alpha_2^F G E_{cb} \\
& + \alpha_3 X_{cb} + \alpha_3^F G X_{cb} + \alpha_4 H_{cbt-1} + \alpha_4^F G H_{cbt-1} + u_{cbt}
\end{aligned} \tag{27}$$

where the superscript  $F$  refers to females and we allow each explanatory variable to have a gender-specific effect on health. Taking gender differences ( $\Delta = \text{females} - \text{males}$ ), we obtain

$$\begin{aligned}
\Delta H_{cbt} = & \alpha_0^F + \alpha_1 \Delta B_{cb(t-1)} + \alpha_1^F B_{cb(t-1)}^F + \alpha_2 \Delta E_{cb} + \alpha_2^F E_{cb}^F + \alpha_3 \Delta X_{cb} + \alpha_3^F X_{cb}^F + \\
& + \alpha_4 \Delta H_{cbt-1} + \alpha_4^F H_{cbt-1}^F + \Delta u_{cbt}
\end{aligned} \tag{28}$$

In this specification,  $\alpha_1$  and  $\alpha_1 + \alpha_1^F$  are the effects of health behaviors lagged once for males and females respectively. Similarly, the gender gap in the "returns" to education is given by coefficient  $\alpha_2^F$ .

Differencing by gender eliminates all unobserved factors that are common to males and females for a given country  $c$ , birth cohort  $b$  and time  $t$ , including genetic and environmental effects, income components, medical inputs and the organization of health care<sup>14</sup>. Even after eliminating common unobservables, however, the residual error component  $\Delta u_{cbt}$  could still be correlated with education and lagged health behaviors. This could happen, for instance, if health conditions and parental background during childhood differ systematically by gender or if labor market discrimination affects individual income and access to health care, conditional on educational attainment. To remove this correlation, we model this residual as

$$\Delta u_{cbt} = \psi_b + \psi_c + \psi_t + \psi_1 \Delta Z_{cbt} + \psi_2 Z_{cbt}^F + \psi_3 \Delta Y_{cbt} + \psi_4 Y_{cbt}^F + v_{cbt} \tag{29}$$

where  $\psi_b$ ,  $\psi_c$  and  $\psi_t$  are cohort, country and year of survey dummies,  $Y$  is real income and  $Z$  is a vector of observables, which includes a rich set of parental background characteristics and health conditions during childhood<sup>15</sup>. Our identifying assumption is that, conditional on these variables which capture gender-specific genetic and en-

<sup>14</sup>See Zweifel and Breyer (1997).

<sup>15</sup>There is a growing literature on the impact of childhood health on adult economic outcomes (Banks et al. (2011) or Smith (2009)). The vector  $Z$  includes: childhood poor health, hospitalization during childhood, presence of serious diseases, had at most 10 books at home at age 10, mother and father in the house at age 10, mother or father died during childhood, number of rooms in the house at age 10, had hot water in the house at age 10, parents drunk or had mental problems at 10, had serious diseases at age 15, born in the country, had proxy interview.

vironmental effects, the error term  $v_{cbit}$  is orthogonal to levels and changes in health behaviors and educational attainment.

For the sake of brevity, we call this method *ADS* (aggregation cum differentiation cum selection on observables). To illustrate, suppose that the key unobservable in (24) is the latent (cell) average ability. The ADS method assumes that part of this latent factor is common across genders and can be differenced out. The residual gender specific component is captured by cohort and country dummies as well as by gender differences in parental background at age 10 and initial health conditions. Conditional on our identification assumption, Eq. (28) is estimated by weighted least squares, using as weight  $\left(\frac{1}{N_M} + \frac{1}{N_F}\right)^{-1}$ , where  $N_M$  and  $N_F$  are the number of males and females in each cell, as suggested by Card and Rothstein (2007).

## 4.2 IV estimates

It is useful to re-write the "reduced form" health equation (20) as follows

$$H_{it} = \omega_1 + \omega_2 E_i + \omega_3 X_{it} + v_{it} \quad (30)$$

where  $i$  is for the individual,  $v$  is the error term and  $Cov(E, v) \neq 0$ . We estimate (30) by instrumental variables, using as instrument for endogenous education the number of years of compulsory education  $YC$ . This is widely considered to be a credible identification strategy, and one that has been extensively used in this and related studies (see Lochner (2011) for a review). We apply this strategy to a multi-country setup, as in Brunello, Fort and Weber (2009), and Brunello, Fabbri and Fort (2009), Fort et al. (2011) and exploit the fact that school reforms have occurred at different points in time in several countries<sup>16</sup>.

For each country and reform included in our sample, we construct pre-treatment and post-treatment samples by identifying for each reform the pivotal birth cohort, defined as the first cohort potentially affected by the change in mandatory school leaving age. We include in the pre- and post-treatment samples all individuals born either before, at the same time or after the pivotal cohort. By construction, the number of years of compulsory education  $YC$  "jumps" with the pivotal cohort and remains at the new level in the post-treatment sample. The timing and intensity of these jumps varies

---

<sup>16</sup>We implement this strategy by selecting 7 countries where the individuals in our sample experienced at least one compulsory school reform: Austria, Denmark, England, France, Italy, the Netherlands and the Czech Republic. The inclusion of the latter country is possible because the IV approach does not require two waves per country. We exclude instead Germany and Sweden because school reforms in these countries were implemented at the regional level and our information on the region where the individual completed her education is not accurate. See Appendix B for a short description of the compulsory school reforms used in this paper.



across countries, and we use the within and across country exogenous variation in the instrument to identify the causal effects of schooling on health.

The vector  $X$  in Eq. (30) includes country fixed effects, cohort fixed effects and country-specific linear or quadratic trends in birth cohorts. These trends account for country-specific improvements in health that are independent of educational attainment<sup>17</sup>. Country fixed effects control for national differences both in reporting styles and in institutions affecting health. Notice that the older cohorts in our data are healthier than average, having survived until relatively old age. Since the comparison of positively selected pre-treatment individuals with younger post-treatment samples is likely to result in a downward bias in the estimates, we control for this selection process by including cohort fixed effects.

## 5 Data

The estimation of the "reduced form" and the dynamic health equation requires data on health outcomes, risky health behaviors, education, parental background and early socio-economic and health conditions. The European Survey of Health, Ageing and Retirement in Europe (SHARE), the English Longitudinal Study of Ageing (ELSA) and their retrospective interviews, satisfy these requirements. SHARE is a longitudinal dataset on the health, socio-economic status and social relations of European individuals aged 50+, and consists of two waves - 2004/5 and 2006/7 - plus a retrospective wave in 2008/9 (SHARELIFE), covering several European countries - Austria, Belgium, Switzerland, Denmark, Spain, France, Germany, Italy, Greece, The Netherlands and Sweden<sup>18</sup>. ELSA has similar characteristics and covers England<sup>19</sup>. Since education is typically accumulated in one's teens or twenties, by focusing on individuals aged 50+ we are considering the long run effects of education on health.

The measure of health used in this paper is self-reported poor health ( $SRPH$ ), a dummy equal to 1 if the individual considers her health as fair or poor and to 0 if she considers it as good, very good and excellent. This is a subjective and comprehensive measure of health, which is conventionally used in the applied literature (Lochner, 2011). One may object that self reported information is likely to be dominated by noise and to fail to capture differences in more objective measures of health.<sup>20</sup> This

---

<sup>17</sup>"Failure to account for secular improvements in health may incorrectly attribute those changes to school reforms, biasing estimates toward finding health benefits of schooling.." (Lochner (2011), p.41)

<sup>18</sup>The Czech Republic, Poland, Israel and Ireland joined in the second wave.

<sup>19</sup>For England, we use waves 2 (2004/5) and 3 (2006/7).

<sup>20</sup>For an early discussion about the importance of measurement error in self-reported health see Bound (1991) and Butler et al. (1987) as well as Baker et al. (2004). These authors were primarily concerned with the impact of measurement error in equations determining the impact of health

is not the case here: among the individuals in the sample who reported poor health, 46% had hypertension, 69% had cardiovascular diseases and 79% suffered some long term illness. On average, they had 2.44 chronic diseases (certified by doctors). In contrast, the percentage of individuals in good health with similar diseases was 28, 44 and 33 percent, respectively.<sup>21</sup> Moreover, the latter group experienced only 1.10 chronic diseases. While our data contain information on chronic diseases, which can be argued to be more objective than self-reported health, we have chosen to focus on the latter in order to be able to compare our results with the bulk of estimates in the relevant literature. However, we also present in the robustness section of this paper some estimates based on the number of chronic diseases<sup>22</sup>.

We measure educational attainment with years of education. The second wave of SHARE provides information on the number of years spent in full time education. In the first wave, however, participants were only asked about their educational qualifications. Thus, for the individuals participating only to the first wave, we calculate their years of schooling using country-specific conversion tables. In ELSA, years of education are computed as the difference between the age when full-time education was completed and the age when education was started. We have four measures of risky health behaviors: whether the individual is currently smoking, whether she drinks alcohol almost every day, whether she refrains from vigorous activity and the body mass index (BMI). These risk behaviors are among the seven listed by the World Health Organization as the most important factors affecting individual health - the remaining three being low fruit and vegetable intake, illicit drugs and unsafe sex.

Table 1 reports the country averages of the health outcome *SRPH*, years of education and annual income (thousand euro at 2005 prices, PPP) in 2006, as well as the means of the four health behaviors (in 2004), separately by gender. There is important cross-country and cross-gender variation, both in the outcome and in health behaviors. As expected, both income and years of education are higher among males aged 50+ than among females of the same age group. The percentage of females reporting poorer health is higher than the percentage of males (32 versus 27 percent). Females are less likely to smoke and drink than males. They have a slightly lower body mass index (26.7 versus 27.1) and tend to exercise vigorously more often than males.

Table A-1 in the Appendix reports the country averages of the parental background variables included in the vector  $Z$ . The table shows that there is important varia-

---

on retirement and other labor market outcomes; justification bias, i.e. non-working persons over-reporting specific conditions is an obvious problem there.

<sup>21</sup>Heiss (2011) finds strong autocorrelation in self-reported health across waves and a strong correlation with future mortality for the HRS.

<sup>22</sup>Using the same dataset, we discuss at length how the education gradient varies with different measures of health in a companion paper (Brunello et al., 2011)

tion both across countries and by gender. For instance, the percentage of individuals with less than 10 books in the house at age 10 ranges from 79% in Italy to 18% in Sweden. The gender gap is particularly relevant in England, where this percentage is 30% for males and 24% for females. Furthermore, the percentage of individuals who was in poor health at age 10 was 9% among Spanish males and 11% among Spanish females. There is less variation between genders in the parental background and housing characteristics: we interpret this as suggestive evidence that parental background characteristics are substantially removed by gender differencing, since within country and cohort they are largely common between males and females, on average.

The estimate of the dynamic health equation (15) requires information on the current and the previous period. The two waves of SHARE and ELSA used in this paper include both individuals who appear in both waves and individuals who are interviewed only in a single wave. We compute cell averages at time  $t$  and  $t - 1$  by using all individuals rather than the longitudinal subsample. Each cell is defined by gender, country, wave and semester of birth. We use semesters rather than years to increase the number of available cells in the estimation<sup>23</sup>, and retain those cells that include at least two observations.

## 6 Results

### 6.1 Baseline Estimates of the Reduced Form and Dynamic Health Equations

As reviewed in Section 2, most earlier contributions to this literature fail to consider the endogeneity of education and health behaviors in their health regressions. For the sake of comparison, we start the illustration of our empirical findings with the estimates of the "reduced form" and dynamic health equations based on micro data. We use a linear probability model and regress self-reported poor health on years of education and a vector of variables, which varies according to whether we consider the "reduced form" or the dynamic health equation but always includes parental and early life controls.

For each regression, we pool males and females but allow for the full set of interactions of each explanatory variable with a gender dummy. Preliminary testing suggests that we cannot reject the null hypothesis that cohort, country, time and early life effects do not vary by gender<sup>24</sup>. We therefore report in Table 2 the results of a more parsimonious specification, which includes a country-specific quadratic trend in

---

<sup>23</sup>Since we do not have information on the month of birth for England, we aggregate by year of birth for this country.

<sup>24</sup>The joint hypothesis is not rejected at the 5 percent level of confidence (p-value: 0.094). We tested separately also the null that the following effects are common between genders: cohort effect (p-value:

the year of birth as well as common cohort, country, year of the survey and early life controls. The table is organized in two columns, one for the "reduced" form equation and the other for the dynamic health equation, which includes health behaviors lagged once, the first lag of health and current income.

Considering first the "reduced form" equation, we find that the marginal effect of one additional year of schooling is equal to  $-0.012$  for males and to  $-0.017$  for females<sup>25</sup>, a relatively small effect when compared to the existing literature for Europe, which points to an effect in the range  $-0.026$  to  $-0.081$  (Lochner (2011), Table 6). This difference could be explained, at least in part, if the education gradient declines with age, given that our sample consists of individuals aged 50+ and the samples used in the literature typically include also younger individuals. Coefficients for parental and early life conditions, including health at age 10, are statistically significant and point in the expected direction: poor health conditions at 10 or 15 as well as poor parental environments at early ages increase self perceived poor health at age 50+. Importantly, the inclusion of these variables reduces the education gradient by 15 to 20 percent<sup>26</sup>, which suggests that they capture at least in part the positive correlation between educational attainment and unobserved individual effects such as ability and initial health.

Turning to the dynamic health equation, we find that our measures of risky health behaviors have statistically significant coefficients, with predictable effects: smoking, refraining from vigorous activity and poor diet leading to higher BMI increase self-perceived poor health. Somewhat unexpectedly, however, drinking alcohol almost every day reduces self-reported poor health. Annual real income also reduces perceived poor health, which exhibits important persistence over time - the lagged dependent variable has a coefficient close to 0.5 but statistically distinct from 1.

Adding health behaviors, income and lagged health reduces the marginal impact of education on health from  $-0.012$  to  $-0.005$  for males, and from  $-0.017$  to  $-0.006$  for females. Assuming that the returns to education for the sample of countries under study is equal to 0.07<sup>27</sup>, the estimated mediating effect of behaviors lagged once is 0.894), country effect (p-value: 0.42), background variables (p-value: 0.263), trends (p-value: 0.112) and we never reject the null at conventional significance levels.

<sup>25</sup>The corresponding semi-elasticities evaluated at the average value of the dependent variable are  $-4.5\%$  for males and  $-5.2\%$  for females. The higher gradient for females could be due to decreasing returns to education, and to the fact that females in our sample are less educated than males. To investigate this point further, we have added to the baseline specification a quadratic term in education but found that it is not statistically significant.

<sup>26</sup>When we exclude parental and early life conditions, the gradient increases in absolute value to 0.016 for males and to 0.020 for females.

<sup>27</sup>See for instance the estimates in Brunello, Fort and Weber (2009). Including income in Eq. (15) implies that  $LRME$  is equal to  $\frac{\pi\sigma_1}{(\pi\sigma_1+\nu+k\rho\bar{Y})}$ , where  $k$  is the coefficient of income in the dynamic health equation,  $\rho$  is the estimated return to education and  $\bar{Y}$  is average income.

9.7% for males and 16.8% for females. In the long run, when we include the effect of earlier health behaviors, the mediating effect almost doubles, to 18.9% for males and 32.3% for females, suggesting that considering only their first lag may substantially under-estimate the contribution of health behaviors to the education gradient. Our estimated long run effects are smaller than those found by Cutler et al. (2008), who use a different approach but conclude that measured health behaviors account for over 40% of the education gradient (on mortality) in a sample of non-elderly Americans<sup>28</sup>.

Although the inclusion of parental and early life controls in our regression is likely to attenuate the correlation between education, health behaviors and unobservables, there is no guarantee that this correlation will disappear entirely. In order to reduce it further, we apply the *ADS* procedure discussed in Section 4.1, which combines aggregation and gender differentiation with selection on observables. The specification tests carried out in the micro data suggest that cohort, country, year and early life effects do not differ significantly by gender. Note that while the value of the cohort, country and year dummies is also common between gender within country and cohorts, the average value of early life variables may differ between genders for a given country and cohort. As a consequence, when we take gender differences of cell data, these common effects are removed together with common unobservables. Therefore, our preferred specification of the *ADS* model includes only differences of early life variables and excludes both country and cohort dummies.<sup>29</sup>

Our results for the *ADS* model are shown on the right-hand side of Table 2, both for the "reduced form" and for the dynamic health equation. When we consider the former, we find that the overall effect of education on health is negative and larger in absolute values for females ( $-0.026$ ) than for males ( $-0.010$ ). Parental and early life variables are jointly statistically significant (p-value: 0.009), mainly because of the gender differences in poor health at age 10. Turning to the dynamic health equation, we find that the effect of education conditional on behaviors is much smaller ( $-0.015$  for females and  $-0.003$  for males). While the precision of the estimates of the effects of behaviors declines with respect to the micro data, we cannot reject the null hypothesis that these effects are jointly statistically significant. Finally, income effects are insignificant and the persistence of self-reported poor health over time is substantially reduced with respect to the estimates based on micro-data.

---

<sup>28</sup>These authors estimate a static health equation, which includes income and occupation among the explanatory variables, and use the following measures of health behaviors: current smoker, ever smoker, number of cigarettes per day, obesity, regular exercise and use of seat belts always.

<sup>29</sup>We also run a less parsimonious specification of the *ADS* model that included cohort and country dummies and tested whether they could be excluded from the model: the null of no cohort and country dummies is never rejected at all conventional level of significance.

Aggregation and differentiation increases the absolute value of the overall education gradient for females from 0.017 to 0.026 but has limited effects on the gradient for males, which marginally declines in absolute value from 0.012 to 0.010. The short and the long run mediating effects of health behaviors are also affected. As shown in Table 4, when compared to the micro-estimates the long run mediating effect for males declines in absolute value (from 0.007 to 0.004) but increases as a share of the gradient (from 18.9 to 44.5%). The opposite happens for females, for whom this effect increases in absolute value from 0.005 to 0.006 but declines as share of the gradient (from 32.3% to 22.8%).

In sum, when we explicitly take into account the endogeneity of education and health behaviors, we find that the mediating effect of the latter ranges between 23 and 45% of the total education gradient. While the effect of education on behaviors does contribute to account for an important share of the gradient, much remains to be explained, either by the role played by unmeasured behaviors or by effects that do not involve behaviors, such as better decision making, stress reduction and more health-conscious peers.

## 6.2 IV Estimates of the Reduced Form Health Equation

In this section, we present the results of our instrumental variables strategy and compare the estimated causal effects of education on health with those obtained from the ADS identification strategy. Our instrument for education is the number of years of compulsory education, which varies across countries and cohorts because of compulsory schooling reforms. For each country, we construct a sample of treated individuals, who have experienced a change in compulsory education, and a control sample, with no change in compulsory schooling. Since our data include only individuals aged 50+, we need to focus on school reforms which took place between the 1940s and the 1960s, and to restrict our attention to a sub-sample of 7 countries affected by these reforms. Table 4 shows the selected countries, the years and the content of the reforms as well as the pivotal cohorts, i.e. the first cohorts potentially affected by the reforms (see Appendix B for a short description of the education reforms used in this paper).<sup>30</sup>

In order to ensure that individuals spent their schooling in their host country, we restrict our sample to individuals aged 50 and above at the time of the interview, who participated in the first or second wave of SHARE (second or third wave in ELSA), and were born in the country or migrated there before age 5. Additionally, we control for

---

<sup>30</sup>We exclude Germany and Sweden because in these countries compulsory schooling laws were implemented gradually across states or municipalities and we cannot identify for the individuals in the sample where their education took place. We use instead data for the Czech Republic (only one wave available), because for the IV-setup we do not need two time periods/waves as in the *ADS* model.

country fixed effects, cohort fixed effects as well as for some individual characteristics (whether the individual is foreign-born, whether there was a proxy respondent for the interview or proxy information is missing and indicators for interview-year). We capture smooth trends in education and health by using country-specific polynomials in cohorts. In particular, we estimate two specifications, one with a linear trend and one with a quadratic trend.

Since the key identifying assumption that changes in individual education can be fully attributed to the reforms is more plausible when the window around the pivotal cohort is small, we estimate our model using two alternative samples, one including individuals who were born up to 10 years before and after the reforms and another where the relevant window is +7,-7. The two samples consist of 15,960 and 12,294 individuals respectively. Table 5 shows the summary statistics by country for the larger sample.

Table 6 shows our estimates of the health-education gradient for both males and females. We report the OLS, 2SLS, ITT (Intention-To-Treat), first stage and IV-Probit estimates for both samples, using two alternative specifications for the country-specific trends (linear or quadratic). The OLS estimate of the gradient is equal to  $-0.017$  for males and to  $-0.025$  for females. When instrumenting years of education with years of mandatory schooling, the magnitude of the coefficients increase somewhat. One additional year of schooling decreases the probability of poor health by 4-8 percentage-points for females and 5-6 percentage-points for males. IV-Probit estimations yield very similar results. The instrumentation strategy works well, our first-stage regressions show that instruments are relevant and not weak (F-Statistics of 13-41): one additional year of compulsory schooling is increasing actual schooling by a quarter to a third of a year. These estimates are comparable with those previously found in the literature using similar identification strategies and represents a plausible reaction, because a compulsory schooling reform should primarily only be effective for individuals who would not have continued schooling in the absence of the reform. We interpret these effects as Local Average Treatment Effects, i.e. the impact of schooling on health for those individuals who were actually affected by the reforms. These individuals typically belong to the lower portion of the education distribution.

### 6.3 IV and ADS Results Compared

Next, we compare the estimated health-education gradient obtained from the IV-approach with the corresponding estimate in the ADS model (Table 7). For the IV-estimates, we report the ones obtained with the linear trend specification for Sample 10. Education decreases self-perceived poor health by 4 and 4.8 percentage points

for females and males respectively. The ADS model yields fairly similar estimates for females (2.6 percentage points), but smaller ones for males (1 percentage point).

Since the IV approach and the ADS-model are based on a different set of countries and cohorts, we re-run the ADS model for the same sample we have used in the IV-approach. The results are shown in the last column of Table 7. The magnitudes increase somewhat to 2.8 percentage points for females and to 2 for males. As mentioned above, while the IV estimates are Local Average Treatment Effects, i.e. the causal effects of education on health for the individuals affected by the compulsory schooling reforms, the estimates obtained from the ADS method pertain to a randomly drawn individual from the entire sample. If the protective effect of education on health is more pronounced for persons with lower education, this could explain the somewhat higher magnitudes in the case of the IV approach.

## 6.4 Robustness Checks

In this sub-section, we discuss several robustness checks. We start by collapsing data by gender, country and year rather than semester of birth. By doing so, we reduce the sample size by almost a half<sup>31</sup>. As shown in the first two columns of Table 8, the effect of education on health is virtually unaffected for females but declines for males. Next, we omit England to take into account that English data are drawn from a different (although quite similar) survey and can only be collapsed by year of birth. The second two columns of Table 8 show that the education gradient changes only marginally. However, when we decompose the gradient into the effect mediated by behaviors and the residual effect, we find that *LRME* in this sub-sample is much smaller than in the full sample, and equal to 8.5% and 11.1% of the gradient for females and males respectively<sup>32</sup>.

Furthermore, we notice that the older cohorts in our data are strongly selected by mortality patterns<sup>33</sup>. To control for this, we add to the regressions the level and the gender difference of life expectancy at birth; these variables vary by country, gender and birth cohort. Since these data are not available for Greece<sup>34</sup>, we are forced to omit that country from the sample. As displayed by the last two columns in the Table, life

---

<sup>31</sup>Recall that for England we do not observe the month of birth. Therefore, cells for England are always aggregated by year of birth.

<sup>32</sup>We have also estimated our equations on two sub-samples of countries, based on their proximity to the Mediterranean Sea, but cannot reject the hypothesis that the estimated coefficients are not statistically different.

<sup>33</sup>Age in our sample ranges from 50 to 86.

<sup>34</sup>We use data on life expectancy at birth from the Human Mortality & Human Life-Table Databases. The databases are provided by the Max Planck Institute for Demographic Research ([www.demogr.mpg.de](http://www.demogr.mpg.de)). The data are missing for some cohorts and for Greece. We use period measures of life expectancy at birth since cohort measures are not available for all the cohorts we considered in the study.



expectancy is never statistically significant in the "reduced form" health equation, and only marginally significant (at the 10% level of confidence) in the dynamic health equation. We conclude that adding this variable does little to our empirical estimates.

We also run our estimates for the sub-sample of individuals aged 50 to 69 and find that one additional year of schooling reduces self-reported poor health by 11.5% for males and by 22.4% for females. These percentages are closer to those found in the empirical literature. Since survivors aged 70 to 86 are typically both better educated and with a stronger protective role of education on health than the average individual in the same age group; i.e. they will have a higher health-education gradient, it is unlikely that the decline of the gradient with age is driven by selection effects. One may think of several alternative reasons for such a decline. For instance, the gradient can fade because cognitive abilities decline with age. On the other hand, the effect of behaviors on health accumulates over time, which increases the gradient. At the same time, one may speculate that differences by education increase with age because the older care more about their health. Our empirical results suggest that the balance of these effects is tilted in favor of the first.

Finally, we consider an alternative and more objective measure of health outcome, the number of chronic diseases. While this number is reported by the interviewed individuals, it is conditional on screening, i.e. each condition must have been detected by a doctor. Table 9 presents both the *ADS* estimates of the "reduced form" and the dynamic health equation, and the *IV* estimates of the "reduced form". Using the *ADS* method, we find evidence of a negative and statistically significant gradient for females ( $-0.057$ ) and of a positive, small and imprecisely estimated gradient for males ( $0.012$ ). The direction of these effects is confirmed but their size in absolute value is larger ( $-0.157$  for females and  $0.080$  for males) when we apply the *IV* method to the sub-sample of 7 countries.

Defining  $P(D)$  as the probability of reporting a condition, this probability is the product of the probability of undergoing screening  $P(S)$  and the probability of having a condition conditional on screening,  $P(D|S)$ . We speculate that in the case of males the positive effect of education on the number of diseases is driven by the fact that better educated males choose more intensive screening. Turning to the decomposition of the gradient into the mediating effect of behaviors and the residual effect, we find that *SRME* and *LRME* for females are equal to 16.5 and 28.1 percent respectively, not far from the effects estimated for self reported poor health. In the case of males, the estimated parameters do not meet the conditions for both *SRME* and *LRME* to be well defined within the range  $[0, 1]$ .

## 7 Conclusions

We propose a strategy to estimate and decompose the health-education gradient which takes into account both the endogeneity of educational attainment as well as the endogenous choice of health behaviors. Our results show that one additional year of schooling reduces self-reported poor health by 7.1% for females and by 3.1% for males. Health behaviors - measured by smoking, drinking, exercising and the body mass index - contribute to explaining this gradient. We find that the mediating effect of behaviors accounts for at most 23% to 45% of the entire effect of education on health, depending on gender. Using a completely different strategy - instrumental variables estimation - we find corroborating results for the health-education gradient.

Since the gradient is key to understanding inequality in health and life expectancy and is also used to assess overall returns to education (Lochner, 2011), it is important to understand the mechanisms governing it. Many of the discussed health behaviors are individual consumption decisions, changes thereof come at personal costs; e.g. abstaining from smoking or drinking good wine. Increases in health achieved by such costly changes in behavior have, thus, to be distinguished from changes resulting from free benefits of education, such as lower stress, better decision making, etc. Moreover, it is relevant for political decisions about subsidizing schooling. If individuals are aware of the health-fostering effects of schooling and these are private, then there is no room for public policy. If individuals are unaware of these benefits, the case for public policy is stronger if health benefits of schooling are primarily free rather than being based on costly health behavior decisions of individuals (Lochner, 2011).

Table 1: Descriptive statistics, baseline estimation sample (micro-data), males (M) and females (F).

Country	Self-rep poor health		Education		Income		Age		Obs	
	M	F	M	F	M	F	M	F	M	F
Austria	0.27	0.31	11.04	9.47	18.74	10.74	65.14	66.18	260	364
Belgium	0.24	0.29	12.36	11.55	16.09	10.82	65.24	65.59	905	1044
Denmark	0.21	0.26	11.25	10.98	16.34	13.02	64.57	65.68	385	399
England	0.28	0.29	11.26	11.20	20.67	14.25	67.50	67.35	1673	2050
France	0.32	0.38	12.17	11.29	23.53	14.04	65.36	66.35	486	638
Germany	0.29	0.35	13.58	12.23	24.50	8.57	65.23	63.69	310	342
Greece	0.19	0.25	9.49	8.16	14.95	6.90	65.10	64.78	717	801
Italy	0.38	0.50	8.08	7.11	13.07	6.55	66.42	65.16	602	722
Netherlands	0.26	0.29	11.88	11.23	22.92	11.29	65.33	64.66	526	599
Spain	0.39	0.52	7.99	7.50	13.65	5.52	67.30	66.44	364	458
Sweden	0.22	0.26	11.42	11.61	16.81	13.00	65.94	65.38	512	615
Switzerland	0.12	0.18	12.25	10.68	29.89	14.10	66.01	64.85	197	232
All	0.27	0.32	11.02	10.37	18.66	11.17	66.03	65.86	6937	8264

Country	Smoking <sub>-1</sub>		Drinking <sub>-1</sub>		No vigorous exercise <sub>-1</sub>		BMI <sub>-1</sub>	
	M	F	M	F	M	F	M	F
Austria	0.21	0.05	0.17	0.17	0.64	0.73	27.46	26.94
Belgium	0.37	0.20	0.20	0.12	0.61	0.75	26.95	26.06
Denmark	0.37	0.20	0.31	0.28	0.48	0.52	26.49	25.57
England	0.22	0.14	0.13	0.12	0.75	0.81	27.81	28.15
France	0.52	0.24	0.19	0.09	0.59	0.73	26.57	25.74
Germany	0.26	0.11	0.21	0.14	0.44	0.43	26.83	26.04
Greece	0.18	0.03	0.36	0.20	0.60	0.67	27.11	26.73
Italy	0.60	0.29	0.25	0.14	0.65	0.74	27.11	26.56
Netherlands	0.38	0.28	0.24	0.24	0.52	0.54	26.26	26.17
Spain	0.45	0.11	0.29	0.10	0.63	0.74	27.62	27.98
Sweden	0.10	0.03	0.12	0.20	0.48	0.60	26.55	25.53
Switzerland	0.34	0.19	0.24	0.19	0.48	0.57	25.78	24.76
All	0.32	0.16	0.21	0.15	0.61	0.70	27.07	26.72

Table 2: Baseline Results - Micro and ADS Model

	Micro-estimates		ADS-model	
	Reduced form	Dynamic HE	Reduced form	Dynamic HE
<b>Females</b>				
education	-0.017 (0.001)***	-0.006 (0.001)***	-0.026 (0.005)***	-0.015 (0.005)***
self-rep poor health <sub>t-1</sub>		0.479 (0.012)***		0.246 (0.046)***
drinking <sub>t-1</sub>		-0.025 (0.012)**		-0.013 (0.053)
smoking <sub>t-1</sub>		0.052 (0.012)***		-0.034 (0.056)
No vigorous exercise <sub>t-1</sub>		0.032 (0.009)***		0.040 (0.042)
BMI <sub>t-1</sub>		0.007 (0.001)***		0.003 (0.004)
income <sub>t</sub>		-0.000 (0.000)**		-0.002 (0.001)
<b>Males</b>				
education	-0.012 (0.001)***	-0.005 (0.001)***	-0.010 (0.005)*	-0.003 (0.005)
self-rep poor health <sub>t-1</sub>		0.486 (0.014)***		0.308 (0.046)***
drinking <sub>t-1</sub>		-0.041 (0.010)***		-0.062 (0.038)
smoking <sub>t-1</sub>		0.030 (0.011)***		0.043 (0.042)
No vigorous exercise <sub>t-1</sub>		0.049 (0.009)***		0.089 (0.041)**
BMI <sub>t-1</sub>		0.006 (0.001)***		0.011 (0.005)**
income <sub>t</sub>		-0.000 (0.000)**		-0.001 (0.001)
<b>Early life</b>				
few books in HH	0.043 (0.009)***	0.022 (0.008)***	0.053 (0.035)	0.040 (0.033)
serious diseases at 15	0.017 (0.008)**	0.004 (0.007)	0.028 (0.036)	0.004 (0.035)
poor health at 10	0.117 (0.014)***	0.062 (0.012)***	0.158 (0.052)***	0.135 (0.049)***
hospital at 10	0.032 (0.016)**	0.025 (0.014)*	0.004 (0.063)	0.042 (0.061)
<i>Principal components</i>				
parents drink or have mental problems at 10	0.036 (0.009)***	0.018 (0.008)**	0.011 (0.039)	0.025 (0.038)
parental absence at 10	0.011 (0.011)	0.007 (0.009)	-0.008 (0.039)	-0.009 (0.037)
poor housing at 10	0.016 (0.004)***	0.013 (0.004)***	0.023 (0.017)	0.014 (0.016)
Cohort effects	yes	yes	no	no
Country-spec trends	yes	yes	no	no
Observations	15,201	15,201	736	734

NOTES: \*\*\*, \*\* and \* indicate statistical significance at the 1-percent, 5-percent and 10-percent level.

Table 3: Decomposition - Micro and ADS Model

	Females		Males	
	Micro-model	ADS-model	Micro-model	ADS-model
Health-Education Gradient (HEG)	-0.017	-0.026	-0.012	-0.010
- behaviors (short-term)	-0.003	-0.004	-0.004	-0.003
- behaviors (long-term)	-0.005	-0.006	-0.007	-0.004
- residual (direct effect)	-0.012	-0.020	-0.010	-0.006
Mediating effect as fraction of HEG				
- SRME (short-term)	0.168	0.172	0.097	0.308
- LRME (long-term)	0.323	0.228	0.189	0.445

NOTES: Computations based on the estimates reported in Table 2.

Table 4: Compulsory schooling reforms in Europe

Country	Reform	Schooling	Pivotal Cohort
Austria	1962/66	8 to 9	1951
Czech Republic	1948	8 to 9	1934
	1953	9 to 8	1939
	1960	8 to 9	1947
Denmark	1958	4 to 7	1947
England	1947	9 to 10	1933
France	1959/67	8 to 10	1953
Italy	1963	5 to 8	1949
Netherlands	1942	7 to 8	1929
	1947	8 to 7	1933
	1950	7 to 9	1936

Table 5: Summary Statistics IV - Sample 10

Country	Self-rep poor health	Education	Compulsory Edu	Age	Obs
Austria	0.233	11.363	8.237	58.971	782
Czech Republic	0.418	12.026	8.535	63.304	2,452
Denmark	0.208	11.802	5.642	59.194	1,898
England	0.373	10.713	9.585	72.355	4,672
France	0.331	11.324	8.275	63.668	2,223
Italy	0.337	8.822	6.032	59.631	2,093
Netherlands	0.338	10.613	8.263	69.95	1,840
All	0.339	10.901	8.088	65.588	15,960

Table 6: Health-Education Gradient - IV approach

	Sample 10		Sample 7	
	lin-trend	qu-trend	lin-trend	qu-trend
<b>Females</b>				
OLS	-0.024 (0.002)***	-0.024 (0.002)***	-0.025 (0.002)***	-0.025 (0.002)***
2SLS	-0.040 (0.024)*	-0.064 (0.034)*	-0.041 (0.035)	-0.085 (0.032)***
ITT	-0.014 (0.008)*	-0.017 (0.008)**	-0.011 (0.009)	-0.023 (0.008)***
First Stage	0.344 (0.053)***	0.253 (0.058)***	0.263 (0.053)***	0.271 (0.058)***
IV-Probit	-0.042 (0.022)*	-0.057 (0.025)**	-0.041 (0.032)	-0.073 (0.017)***
F-Stat (First Stage)	41.93	18.95	24.89	21.66
Observations	8,602	8,602	6,631	6,631
<b>Males</b>				
OLS	-0.017 (0.002)***	-0.017 (0.002)***	-0.017 (0.002)***	-0.017 (0.002)***
2SLS	-0.048 (0.029)*	-0.054 (0.029)*	-0.062 (0.029)**	-0.064 (0.034)*
ITT	-0.016 (0.009)*	-0.018 (0.008)**	-0.020 (0.008)**	-0.020 (0.010)**
First Stage	0.323 (0.076)***	0.318 (0.078)***	0.313 (0.079)***	0.298 (0.082)***
IV-Probit	-0.047 (0.024)**	-0.051 (0.022)**	-0.056 (0.019)***	-0.057 (0.022)***
F-Stat (First Stage)	17.87	16.62	15.66	13.07
Observations	7,358	7,358	5,663	5,663

NOTES: \*\*\*, \*\* and \* indicate statistical significance at the 1-percent, 5-percent and 10-percent level.

Table 7: Health-Education Gradient - IV and ADS compared

	IV-estimate	ADS-model	
		All countries	IV-sample
Females	-0.040 (0.024)*	-0.026 (0.005)***	-0.028 (0.007)***
Males	-0.048 (0.029)*	-0.010 (0.005)*	-0.020 (0.008)**

NOTES: \*\*\*, \*\* and \* indicate statistical significance at the 1-percent, 5-percent and 10-percent level.

Table 8: Robustness - ADS Model

	ADS year panel		ADS w/o ENG		ADS l-exp, w/o GRC	
	Red form	Dynamic HE	Red form	Dynamic HE	Red form	Dynamic HE
<b>Females</b>						
education	-0.025 (0.006)***	-0.011 (0.007)	-0.023 (0.005)***	-0.016 (0.006)***	-0.03 (0.006)***	-0.018 (0.006)***
self-rep poor health <sub>t-1</sub>		0.307 (0.063)***		0.240 (0.046)***		0.252 (0.052)***
drinking <sub>t-1</sub>		0.017 (0.069)		-0.017 (0.052)		-0.031 (0.056)
smoking <sub>t-1</sub>		-0.080 (0.076)		-0.043 (0.056)		-0.031 (0.063)
No vigorous exercise <sub>t-1</sub>		-0.016 (0.057)		0.021 (0.044)		0.036 (0.045)
BMI <sub>t-1</sub>		0.001 (0.005)		0.000 (0.005)		0.002 (0.004)
income <sub>t</sub>		-0.001 (0.002)		-0.003 (0.002)*		-0.003 (0.002)*
<b>Males</b>						
education	-0.006 (0.007)	0.004 (0.007)	-0.008 (0.005)	-0.004 (0.005)	-0.010 (0.006)*	-0.004 (0.006)
self-rep poor health <sub>t-1</sub>		0.301 (0.060)***		0.319 (0.046)***		0.295 (0.051)***
drinking <sub>t-1</sub>		-0.011 (0.051)		0.078 (0.038)**		-0.067 (0.042)
smoking <sub>t-1</sub>		0.001 (0.056)		-0.038 (0.042)		0.038 (0.049)
No vigorous exercise <sub>t-1</sub>		0.076 (0.054)		0.090 (0.043)**		0.077 (0.044)*
BMI <sub>t-1</sub>		0.005 (0.007)		0.014 (0.006)**		0.011 (0.006)**
income <sub>t</sub>		-0.002 (0.001)		-0.001 (0.001)		-0.001 (0.001)
<b>Early life</b>						
few books in HH	0.024 (0.048)	-0.006 (0.047)	0.050 (0.035)	0.051 (0.034)	0.085 (0.038)**	0.076 (0.036)**
serious diseases at 15	0.110 (0.051)**	0.070 (0.050)	0.021 (0.037)	0.007 (0.035)	0.021 (0.038)	-0.006 (0.037)
poor health at 10	0.185 (0.073)**	0.170 (0.070)**	0.137 (0.053)***	0.109 (0.050)**	0.164 (0.053)***	0.146 (0.051)***
hospital at 10	-0.078 (0.093)	-0.028 (0.091)	0.060 (0.065)	0.097 (0.062)	-0.009 (0.065)	0.016 (0.062)
<i>Principal components</i>						
parents drink or have mental problems at 10	-0.015 (0.054)	0.010 (0.053)	0.029 (0.041)	0.043 (0.039)	-0.009 (0.041)	0.011 (0.040)
parental absence at 10	0.047 (0.056)	0.029 (0.054)	-0.022 (0.040)	-0.016 (0.038)	0.009 (0.041)	0.005 (0.039)
poor housing at 10	0.039 (0.023)*	0.029 (0.022)	0.022 (0.017)	0.010 (0.016)	0.014 (0.018)	0.004 (0.018)
<b>Life-expectancy</b>						
females					0.007 (0.005)	0.009 (0.005)*
males					0.005 (0.003)	0.007 (0.004)*
Observations	389	387	701	701	640	638

NOTES: \*\*\*, \*\* and \* indicate statistical significance at the 1-percent, 5-percent and 10-percent level.

Table 9: Number of Chronic Diseases

	ADS	ADS	IV
	Reduced form	Dynamic HE	Sample 10, lin-trend
<b>Females</b>			
education	-0.057 (0.015)***	-0.024 (0.016)	-0.157 (0.091)*
# chronic diseases <sub>t-1</sub>		0.413 (0.044)***	
drinking <sub>t-1</sub>		-0.044 (0.161)	
smoking <sub>t-1</sub>		0.007 (0.178)	
no vigorous exercise <sub>t-1</sub>		0.279 (0.131)***	
BMI <sub>t-1</sub>		0.012 (0.305)	
income <sub>t</sub>		-0.002 (0.004)	
<b>Males</b>			
education	0.012 (0.017)	-0.006 (0.016)	0.080 (0.066)
# chronic diseases <sub>t-1</sub>		0.337 (0.046)***	
drinking <sub>t-1</sub>		-0.089 (0.116)	
smokeing <sub>t-1</sub>		0.045 (0.147)	
no vigorous exercise <sub>t-1</sub>		0.220 (0.198)	
BMI <sub>t-1</sub>		0.041 (0.016)*	
income <sub>t</sub>		-0.004 (0.005)	
<b>Early Life</b>			
few books in HH	-0.135 (0.110)	-0.133 (0.102)	
serious diseases at 15	0.067 (0.114)	0.084 (0.106)	
poor health at 10	0.084 (0.164)	-0.004 (0.151)	
hospital at 10	0.081 (0.200)	0.112 (0.186)	
<i>Principal components</i>			
parents drink or have mental prbs	0.149 (0.124)	0.124 (0.117)	
parental absence at 10	-0.128 (0.123)	-0.112 (0.114)	
poor housing at 10	0.069 (0.054)	0.037 (0.050)	



## References

- Adams, Scott J. (2002), 'Educational Attainment and Health: Evidence from a Sample of Older Adults', *Education Economics* **10**(1), 97–109.
- Albouy, V. and L. Lequien (2009), 'Does Compulsory Education Lower Mortality?', *Journal of Health Economics* **28**(1), 155–168.
- Arendt, J.N. (2005), 'Does Education Cause Better Health? a Panel Data Analysis using School Reforms for Identification', *Economics of Education Review* **24**(2), 149–160.
- Arendt, J.N. (2008), 'In Sickness and in Health - Till Education Do Us Part: Education Effects on Hospitalization', *Economics of Education Review* **27**(2), 161–172.
- Baker, Michael, Mark Stabile and Catherine Deri (2004), 'What do self-reported, objective, measures of health measure?', *Journal of Human Resources* **39**(4), 1067–1093.
- Banks, James, Zoe Oldfield and James P. Smith (2011), Childhood health and differences in late-life health outcomes between England and the United States. NBER WP 17096.
- Bound, John (1991), 'Self-reported versus objective measures of health in retirement models', *Journal of Human Resources* **26**(1), 106–138.
- Braakmann, Nils (2011), 'The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England', *Journal of Health Economics* **30**, 753–763.
- Brunello, Giorgio, Daniele Fabbri and Margherita Fort (2009), Years of schooling, human capital and the body mass index of European females, IZA Discussion Papers 4667, Institute for the Study of Labor (IZA).
- Brunello, Giorgio, Margherita Fort and Guglielmo Weber (2009), 'Changes in compulsory schooling, education and the distribution of wages in Europe', *Economic Journal* **119**(March), 516–539.
- Brunello, Giorgio, Margherita Fort, Nicole Schneeweis and Rudolf Winter-Ebmer (2011), The causal effect of education on health: Evidence from older Europeans. mimeo, University of Linz, Austria.

- Butler, J.S., Richard Burkhauser, Jean M. Mitchell and Theodore P. Pincus (1987), ‘Measurement error in self-reported health variables’, *The Review of Economics and Statistics* **69**(4), 644–650.
- Card, D. and J. Rothstein (2007), ‘Racial Segregation and the Black-White Test Score Gap’, *Journal of Public Economics* **91**, 2158–2184.
- Cawley, John and Christopher Ruhm (2011), The Economics of Risky Health Behaviors, Working Paper 17081, NBER.
- Clark, D. and H. Royer (2010), The Effect of Education on Adult Health and Mortality: Evidence from Britain, Working Paper 16013, National Bureau of Economic Research.
- Conti, G., J. Heckman and S. Urzua (2010), ‘The Education-Health Gradient’, *American Economic Review: Papers and Proceedings* **100**, 234–238.
- Contoyannis, P. and A.M. Jones (2004), ‘Socio-economic status, health and lifestyle’, *Journal of Health Economics* **23**, 965–995.
- Cutler, D.M. and A. Lleras-Muney (2006), Education and Health: Evaluating Theories and Evidence, Working Paper 12352, National Bureau of Economic Research.
- Cutler, D.M. and A. Lleras-Muney (2010), ‘Understanding Differences in Health Behaviour by Education’, *Journal of Health Economics* **29**, 1–28.
- Cutler, D.M., A. Lleras-Muney and T. Vogl (2008), Socioeconomic Status and Health: Dimensions and Mechanisms, Working Paper 14333, National Bureau of Economic Research.
- Cutler, D.M., E. L. Glaeser and J. M. Shapiro (2003), ‘Why Have Americans Become More Obese?’, *Journal of Economic Perspectives* **17**(3), 93–118.
- Feinstein, Leon, Ricardo Sabates, Tashweka Anderson, Annik Sorhaindo and Cathie Hammond (2006), *Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium*, Paris, chapter What are the effects of education on health?, pp. 171–354.
- Finkelstein, A., E.F.P. Luttmer and M.J. Notowidigdo (2008), What Good Is Wealth Without Health? the Effect of Health on the Marginal Utility of Consumption, Working Paper 14089, National Bureau of Economic Research.

- Fort, M., N. Schneeweis and R. Winter-Ebmer (2011), More Schooling, More Children: Compulsory Schooling Reforms and Fertility in Europe, Working Paper 1105, Department of Economics, Johannes Kepler University of Linz.
- Fort, Margherita (2006), 'Education reforms across Europe: A toolbox for empirical research'. Paper version: May 11, 2006, mimeo.
- Garrouste, Christelle (2010), *100 years of educational reforms in Europe: A contextual database*, European Commission Joint Research Center, Luxembourg: Publications Office of the European Union.
- Grossman, M. (1972), 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy* **80**, 223–255.
- Heiss, Florian (2011), 'Dynamics of self-rated health and selective mortality', *Empirical Economics* **40**, 119–140.
- Juerges, H., E. Kruk and S. Reinhold (2009), The Effect of Compulsory Schooling on Health: Evidence from Biomarkers, Discussion Paper 183–2009, Mannheim Institute for the Economics of Aging (MEA).
- Kempton, D., H. Juerges and S. Reinhold (2011), 'Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany', *Journal of Health Economics* **30**(2), 340–354.
- Lochner, L. (2011), Non-Production Benefits of Education: Crime, Health, and Good Citizenship, Working Paper 16722, National Bureau of Economic Research.
- Mazumder, B. (2008), 'Does Education Improve Health: A Reexamination of the Evidence from Compulsory Schooling Laws', *Economic Perspectives* **33**(2), 1–15.
- Meghir, Costas, Marten Palme and Emilia Simeonova (2011), Education, health and mortality: Evidence from a social experiment. mimeo, Stockholm University.
- Oreopolous, Phillip (2007), 'Do Dropouts Drop Out Too Soon? Wealth, Health, and Happiness from Compulsory Schooling', *Journal of Public Economics* **91**(11–12), 2213–2229.
- Park, C. and C. Kang (2008), 'Does Education Induce Healthy Lifestyle?', *Journal of Health Economics* **27**(6), 1516–1531.
- Powdthavee, N. (2010), 'Does Education Reduce the Risk of Hypertension? Estimating the Biomarker Effect of Compulsory Schooling in England', *Journal of Human Capital* **4**(2).

- Rosenzweig, M.R. and T. P. Schultz (1983), 'Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight', *Journal of Political Economy* **91**(5), 723–746.
- Ross, C.E. and C. Wu (1995), 'The Links Between Education and Health', *American Sociological Review* **60**(5), 719–745.
- Silles, M.A. (2009), 'The Causal Effect of Education on Health: Evidence from the United Kingdom', *Economics of Education Review* **28**(1), 122–128.
- Smith, James P. (2009), 'The Impact of Childhood Health on Adult Labor Market Outcomes', *The Review of Economics and Statistics* **91**(3), 478–489.
- Stowasser, Till, Florian Heiss, Daniel McFadden and Joachim Winter (2011), "healthy, wealthy and wise?" revisited: An analysis of the causal pathways from socio-economic status to health. NBER WP 17273.
- Zweifel, Peter and Friedrich Breyer (1997), *Health Economics*, Oxford University Press, Oxford.

## A Synthetic Indicators for Parental Background

We have built synthetic indicators for parental background variables in order to reduce the dimensionality of the vector of controls by extracting the first principal component from several group of indicators. Since most indicators are discrete we use the polychoric or polyserial correlation matrix instead of the usual correlation matrix as the starting point of the principal component analysis. The polychoric correlation matrix is a maximum likelihood estimate of the correlation between ordinal variables which uses the assumption that ordinal variables are observed indicators of latent and normally distributed variables. The polyserial correlation matrix is defined in a similar manner when one of the indicator is ordinal and the others are continuous. We list below the synthetic indicators, the observed variables used for each indicator and the interpretation we propose, based on the sign of the scoring coefficients. The scoring coefficients are the same across males and females (otherwise, we argue, results would not be comparable and we could not proceed with the aggregation-differentiation strategy).

**Housing at 10** based on the number of rooms in the house at age 10 and facilities in the house (hot water) at age 10. The extracted first principal component decreases as the number of rooms in the house (where the individual lived at age 10) increases and if there was no hot water: we interpret this indicator as *poor housing conditions at age 10*;

**Parental background at 10** based on binary indicators of whether parents drunk or had mental problems when the individual was aged 10. Since the extracted principal component increases if parents drunk or had mental problems, we interpret it as *poor parental background at age 10*;

**Parental absence/presence at 10** based on three binary indicators: whether the mother died early, whether the father died early and whether the mother and the father were present when the individual was aged 10. The extracted principal component increases if any parent died early and decreases when parents were present at age 10. We interpret this indicator as *poor care at young age*.

Descriptive statistics on the background variables used to build the synthetic indicators and the additional background variables used in the baseline specification are reported in Table A-1.

Table A-1: Descriptive statistics, baseline estimation sample (micro-data), males (M) and females (F).

Country	Few books at 10		Serious dis. at 15		Poor Health at 10		Hospital at 10		No hot water at 10		Rooms at 10	
	M	F	M	F	M	F	M	F	M	F	M	F
Austria	0.42	0.48	0.33	0.32	0.13	0.13	0.11	0.10	0.37	0.37	3.3	3.1
Belgium	0.49	0.46	0.27	0.28	0.06	0.09	0.04	0.05	0.30	0.33	5.1	5.2
Denmark	0.23	0.24	0.25	0.25	0.08	0.08	0.09	0.09	0.13	0.14	4.4	4.3
England	0.30	0.24	0.36	0.31	0.10	0.13	0.11	0.11	0.04	0.21	2.9	3.0
France	0.47	0.48	0.29	0.28	0.10	0.13	0.04	0.04	0.24	0.26	4.3	4.0
Germany	0.32	0.31	0.30	0.33	0.13	0.12	0.09	0.08	0.10	0.10	3.9	4.0
Greece	0.64	0.64	0.21	0.17	0.00	0.00	0.00	0.01	0.38	0.33	2.7	2.8
Italy	0.79	0.75	0.16	0.21	0.05	0.08	0.02	0.03	0.47	0.45	3.1	2.9
Netherlands	0.35	0.30	0.23	0.22	0.11	0.11	0.08	0.08	0.05	0.04	4.7	4.6
Spain	0.66	0.65	0.14	0.17	0.09	0.11	0.02	0.02	0.46	0.44	3.6	3.5
Sweden	0.20	0.18	0.24	0.24	0.06	0.08	0.09	0.08	0.14	0.13	3.7	3.6
Switzerland	0.28	0.31	0.30	0.32	0.06	0.14	0.07	0.07	0.03	0.05	4.8	4.9
All	0.43	0.41	0.27	F	0.08	F	0.07	0.07	0.21	0.21	3.7	3.7

Country	Parents drunk at 10		Parents ment. prob. at 10		Moth/Fath present at 10		Mother died early		Father died early	
	M	F	M	F	M	F	M	F	M	F
Austria	0.09	0.09	0.02	0.02	0.80	0.71	0.0	0.0	0.0	0.0
Belgium	0.09	0.09	0.01	0.03	0.92	0.92	0.0	0.0	0.0	0.0
Denmark	0.07	0.09	0.08	0.09	0.89	0.90	0.0	0.0	0.0	0.0
England	0.05	0.06	0.05	0.06	0.89	0.89	0.01	0.0	0.01	0.01
France	0.10	0.10	0.01	0.01	0.90	0.86	0.01	0.01	0.01	0.01
Germany	0.07	0.08	0.04	0.05	0.79	0.84	0.0	0.0	0.0	0.0
Greece	0.05	0.05	0.00	0.00	0.97	0.97	0.0	0.0	0.0	0.0
Italy	0.10	0.11	0.01	0.00	0.92	0.93	0.0	0.0	0.0	0.0
Netherlands	0.02	0.05	0.02	0.03	0.92	0.92	0.0	0.0	0.0	0.0
Spain	0.08	0.07	0.01	0.01	0.87	0.88	0.0	0.0	0.0	0.0
Sweden	0.07	0.08	0.02	0.02	0.87	0.88	0.0	0.0	0.0	0.0
Switzerland	0.09	0.09	0.03	0.03	0.91	0.94	0.0	0.0	0.0	0.0
All	0.07	0.08	0.03	0.03	0.90	0.90	0.0	0.0	0.01	0.0

## B Educational Reforms in Europe

In this section, we briefly describe the compulsory schooling reforms we are using in this study. The choice of reforms differs somewhat from Brunello, Fabbri and Fort (2009) and Brunello, Fort and Weber (2009) because the individuals in our data are aged 50 or older at the time of the interviews in 2004/2006. Therefore, we need to focus only on relatively early reforms. For further details on educational reforms in Europe see Fort (2006).

**Austria** In 1962 a federal act was passed that increased compulsory schooling from 8 to 9 years. The law came into effect on September 1, 1966. Pupils who were 14 years old (or younger) at that time had to attend school for an additional year. Since compulsory education starts at the age of 6 and the cut-off date for school-entry is September 1, (mostly) individuals born between September and December 1951 were the first ones affected by the reform. Thus, the pivotal cohort is 1951.

**Czech Republic** In the 20<sup>th</sup> century, compulsory education was reformed several times. In 1948 compulsory schooling was increased from 8 to 9 years (age 6 to 15). It was reduced to 8 in 1953 and increased to 9 again in 1960. Two further changes took place in 1979 and 1990. We consider all three reforms for our analysis. The pivotal cohorts are 1934 (for the first reform), 1939 (for the second) and 1947 for the reform in 1960. See Garrouste (2010) for more information on compulsory schooling reforms in the Czech Republic.

**Denmark** In 1958 compulsory education was increased by 3 years, from 4 to 7. In 1971 compulsory schooling was further increased by 2 years, from 7 to 9. Education started at age 7, thus pupils who were 11 years old (or younger) in 1958 were potentially affected by the first reform, i.e. children born in 1947 and after. Since our data only cover individuals 50+ in 2004/2006, we only consider the first reform for this study.

**England** Two major compulsory schooling reforms were implemented in the UK in 1947 and 1973. The first reform increased the minimum school leaving age from 14 to 15, the second reform from 15 to 16. Since the school-entry age is 5 in the UK, compulsory schooling was increased from 9 to 10 years in 1947 and from 10 to 11 years in 1973. Pupils who were 14 years old (or younger) in 1947 were affected by the first reform, i.e. cohorts born in 1933 and after. Due to the sampling frame of ELSA (individuals 50+), we only consider the first reform in this study.

**France** Two education reforms were implemented in France. In 1936, compulsory schooling was increased from 7 to 8 years (age 13 to 14) and in 1959 from 8 to 10 years

(age 14 to 16). After a long transition period, the second reform came into effect in 1967. The first reform affected pupils born 1923 (and after) and the second reform pupils born 1953 (and after).

**Italy** In 1963 junior high school became mandatory in Italy and compulsory years of schooling increased by 3 years (from 5 to 8 years). The first cohort potentially affected by this reform is the cohort born in 1949.

**Netherlands** The Netherlands experienced many changes in compulsory education in the last century. In this paper, we consider three education reforms: in 1942, in 1947 and in 1950. With the first reform compulsory schooling was increased from 7 to 8 years, with the second reform it fell back to 7 years and with the last reform it increased again by 2 years, from 7 to 9. Accordingly, we choose the cohorts born in 1929, 1933 and 1936 as pivotal cohorts.