

Secondary School Fees and the Causal Effect of Schooling on Health Behavior

Steffen Reinhold Hendrik Jürges

MEA – University of Mannheim
L13,17
68131 Mannheim
Germany

April 2009

Abstract

Using German census data we estimate the causal effect of education on smoking and overweight/obesity using the abolition of secondary school fees as instrumental variable. The West German federal states enacted this reform at different dates after World War II generating exogenous variation in the access to secondary education. While we find a strong association between schooling and health behaviors using OLS, we do not find support for the notion that education causes better health behavior.

Keywords: Education, Health, Health Behaviors, School Fees, Instrumental Variables

JEL-Codes: I12, I21

1 Introduction

There is an almost ubiquitous association between education and health (see Grossman (2006) and Cutler and Lleras-Muney (2006) for reviews of the economic literature on the education-health gradient). However, it is still unclear how much of this association is due to a causal relationship running from education to health. If there is a causal relationship another important question would be how education improves health. Much of the recent economic literature exploits exogenous variations in education policy, particularly in compulsory school leaving age, as natural experiments to estimate the causal effect of education on health using instrumental variable or regression discontinuity techniques. The evidence emerging from this literature is mixed Lleras-Muney (2005), Oreopolous (2006), and Silles (2009) find strong positive effects of education on mortality and self-rated health. In contrast, Clark and Royer (2008) and Albouy and Lequien (2008) find hardly any effect of education on health outcomes. Even if education improves health outcomes, the actual mechanism is still unclear.. One possibility is that better educated make greater investments in their health. There is now a large literature investigating whether education improves health behaviors, in particular smoking and overweight as important correlates of health outcomes. Sander (1995) finds evidence that better educated are more likely to quit smoking using family background as instrumental variables.. Currie and Moretti (2003) use the college openings as instruments for maternal education. While their focus is on child health they also find that education reduces the likelihood that the mother smokes. Similarly, Walque (2007) and Grimard and Parent (2007) find that college education reduces smoking rates using draft avoidance strategies during the Vietnam war as instruments. While there is some evidence that education reduces smoking rates the picture is less clear for overweight and obesity. For instance, Kenkel et al. (2006) find that high school completion in the US reduces smoking rates but the effect on

overweight is much smaller. Similarly, Arendt (2005) uses schooling reforms in Denmark as instruments. His results suggest that education also reduces overweight, but the estimates are too imprecise to reject the null hypothesis of no effect.

When interpreting results from IV estimates some caution is required. If the effect of schooling is heterogeneous across the population IV estimates only identify the causal effect for individuals who are affected by the reforms (Imbens and Angrist, 1994). Thus, when using compulsory schooling laws the estimated effects are valid only for individuals with low educational attainment because these laws are not binding for persons with university education. Thus, it is important to consider different sources of exogenous variation as instruments to assess the effect of education on health across the whole distribution.

We contribute to this literature by using the abolition of secondary school fees in Germany in the 1950s as a source of exogenous variation affecting a different sub-population of individuals than compulsory schooling reforms. The post war West German secondary school system was characterized by three different ability-related tracks, the *basic*, *intermediate* and *academic* track, each leading to a specific leaving certificate. The basic and intermediate tracks prepare for further vocational training whereas the academic track prepares for later studies at the university. After fourth grade in elementary school, students had the opportunity to stay in the free basic track (up to grade 8 or 9) or switch to either middle school (intermediate track up to grade 10) or grammar school (academic track up to grade 13). Upper secondary education (grades 11 to 13) was offered in the academic track only. Figure 1 presents an overview of the different tracks in the German educational system.

School fees had to be paid in both middle schools and grammar schools. Switching track after the initial decision was theoretically possible, but such instances were rare because the tracks

differ not only in their years of schooling but also in their curricula. Thus, at about age 10, the students or their parents had to decide on the secondary school track and consequently on the years of additional schooling. Only graduates of the academic track became eligible to attend college or university.

Fees for the academic track in grades 5 to 13 were substantial, typically amounting to about ten percent of an average worker's gross earnings per pupil, while fees in the intermediate track were typically lower. After World War II, education policy became a state responsibility. School fees were abolished state by state between 1947 (Bremen) and 1962 (Rhineland-Palatinate) for all tracks; for details see Riphahn (2008), Figures 3 and 4. It is difficult to ascertain a systematic pattern in fee abolition: Riphahn (2008) estimates hazard models for fee abolition based on state-level panel data. She finds that abolition of fees is not statistically significantly related to indicators for social-democratic governments, per capita income, or income growth rates. For our identification strategy we need to assume that fee abolition only affects secondary schooling decisions but has no direct effects on health. This assumption would be violated if for example states where people have political preferences for supporting education also prefer higher investments in health. Similarly, if states which grow faster invest more in education and health, then the exclusion restriction could be violated. Riphahn's (2008) results give some confidence in the identifying strategy as fee abolition does not seem to be related to political preferences or per capita income and growth.

In some states like Lower Saxony, the fees for the intermediate track were abolished a couple of years before the fees for the academic track, but unfortunately we do not have this detailed information for all German states. The abolition of secondary school fees has decreased the relative costs of attending the intermediate and academic track. Riphahn (2008) shows that the abolition of secondary school fees had a modest positive impact on the probability of

attending the academic track and thus on the number of years in school. This increase in years of schooling can also be seen below in our discussion of the first stage results of the instruments.

2 Data Description

We use data from the German Microcensuses of 1999 and 2003 to analyze smoking behavior, weight problems, and vaccination behavior. The Microcensus is an annual administrative sample survey covering 1 percent of the German population. For several reasons, we can only use part of the original sample for our analysis. First, we exclude all individuals younger than 21 because secondary schooling might not be completed at the time of the survey. Second, we exclude all individuals older than 65 because health related questions were asked of working-age individuals only. Furthermore, we lose more than half of the 21-65 year olds because health-related questions were given to a 45% random subsample of the Microcensus. Although participation in the Microcensus is compulsory, answering questions on education and health questions is (partly) voluntary, so that our sample is further reduced due to item non-response. The overall sample size used in this paper is still sizeable and amounts to roughly 200,000 individuals.

Our measures of health behavior are related to smoking and weight problems. Using information on whether individuals are current or former smokers, we construct indicator variables for current smokers, ever smokers, and quitting smoking. Based on self-reported weight and height, we construct indicator variables for being overweight ($BMI > 25$) and obese ($BMI > 30$). Summary statistics for health behavior can be found in Table 1. Roughly one third of the sample are current smokers and about one half have ever smoked, with men being more likely to smoke currently and to have ever smoked than women. Quitting rates are somewhat higher than one third and similar for men and women. Almost one half of the sample are

overweight and one tenth is obese. Again, we find some gender differences, with men being more often overweight and obese.

Years of schooling are not directly measured in the Microcensus. Instead, education is measured in levels, as completed secondary school tracks. Similar to Pischke and von Wachter (2008) we recode this qualitative information into years of schooling using the number of years required to complete each track (note that we do not include years of tertiary education). In addition, we use information on year of birth and compulsory schooling laws in the state of residence to infer whether individuals who attended the basic track have completed eight or nine years of schooling. We use years of schooling as education variable because our instrument affects the costs of attending both the intermediate and the academic track. Each track offers additional schooling beyond the mandatory level (either 8 or 9 years). Table 1 shows that the average number of years in primary and secondary school is about 10, with men having attended .16 years more than women.

We also considered a model using completion of the academic track as treatment variable, but we found that the school fees had only little explanatory power for the decision to attend the academic track, possibly because students at the margin of attending the academic track had to pay fees in the intermediate track as well. Thus, abolition of the fees in both tracks did not change their behavior much. School fees may have played a much larger role for students at the margin of attending the intermediate track. These students had to decide whether they stay in the free basic track for 4 or 5 years of additional schooling or switch to the intermediate track with 6 additional years where they had to pay fees for all 6 years.

Our instrument is a dummy variable that indicates whether an individual ever had to pay school fees to attend the tracks beyond basic education ("fees"). More precisely, this variable

indicates whether fees had to be paid by the time when the individual was 10 years old. Since fees were not re-introduced, individuals who did not have to pay fees at this age never had to pay fees throughout their educational careers. Similar to Riphahn (2008), we also consider a second instrument for transitional periods (“transition”). This transitional period starts 3 to 4 years before the complete abolition. Since school fees for the intermediate track may have been abolished some time before the abolition of fees for the academic track this may be the more relevant instrument for years of schooling for students at the margin of attending the intermediate track versus the basic track. As a robustness check we present results using this other definition leading to qualitatively very similar results.

3 Empirical Specification and Results

We estimate linear probability models by OLS and IV. Smoking behavior and weight problems are modeled as a function of state fixed effects, cohort fixed effects, and years of education (in primary and secondary school). Angrist (2001) suggests estimating linear probability models using 2SLS even in the context of limited dependent variables when the parameter of interest are causal effects rather than the structural parameters of the underlying latent indices.

3.1 The Association between Educational Attainment and Health Behavior

Our OLS model can be written in the form:

$$H_i = \beta_0 + \beta_1 EducYears + \eta_{State} + \mu_{Cohort} + \lambda_{Year} + \delta_{Sex} + \varepsilon_i \quad (1)$$

where H is an indicator variable for health behavior. We use a full set of state of residence dummies, and year of birth dummies to control non-parametrically for any state and cohort effects on health behavior. We also include a dummy variable indicating the survey year and a

sex dummy when we use the pooled sample of men and women. Table 2 shows our OLS regression results. Consistent with much of the previous empirical research we find a strong relationship between education and health behavior. Each additional year in school is associated with a decrease in current smoking rates by around four percentage points, and the association is somewhat stronger for men than for women. The association with ever having smoked is of similar magnitude showing that one additional year of schooling is associated with a reduction of the probability that the individual ever smoked by around three percentage points. Education is also positively associated with the probability that the individual stops smoking. Each additional year is associated with an increase in this probability of around three percentage points, and this relationship is stronger for women than for men. Further, education is also associated with lower rates of weight problems both for men and women. Overall, our results are consistent with much previous research showing a strong correlation between higher educational attainment and better health behavior.

3.2 The Causal Effect of Educational Attainment on Health Behavior

The OLS coefficient on years of primary and secondary schooling is potentially biased mainly because of self-selection into more education based on unobservable characteristics. Instrumental variables are one standard way of approaching this problem. The instruments have to satisfy two conditions. They must be exogenous and relevant. The first condition states that they have an influence on the outcome of interest – conditional on covariates – only through their effect on the endogenous regressor. With one instrument, this assumption is inherently untestable. However, we believe that controlling in a very flexible way for any cohort effects and state effects by including state of residence and cohort fixed effects lends credibility to our exogeneity assumption.

The second assumption (relevance) can be assessed empirically. We present F-tests for the exclusion of the instrument in the first stage regression in Table 3. In all equations the same dummy for no school fees at age 10 is used. Thus, the same equation is estimated in all columns. However, the sample sizes differ because the outcome variables are sometimes missing. In addition, the likelihood of quitting to smoke can only be estimated on the sample of former and current smokers. As a rule of thumb (Staiger and Stock, 1997), F-statistics should be above 10. Values below this threshold would indicate a weak instruments problem. Since we use different subsamples of varying size, we show first stage F-statistics for each of them. Consistent with Riphahn (2008) we find a small but statistically significant effect of secondary school fees on years of schooling in primary and secondary schooling. On average, the presence of school fees reduced the number of years of schooling by around 0.1 years. Only when we restrict the sample to men who have ever smoked in their lives the instrument becomes weak.

In Table 4 we present IV regression results for the effect of education on health behavior using school fees as instruments for years of schooling. For smoking behavior we find no significant causal effect no matter whether we look at current smoking or ever smoking. Higher educational attainment also has no statistically significant effect on quitting behavior. These findings are in stark contrast to the OLS results where education is associated with lower smoking rates and increased likelihood that individuals quit smoking. Regarding weight problems, we find that the coefficient on education becomes positive but statistically insignificant. Thus, our estimates do not lend support to the notion that education has a positive effect on health behavior. These results are robust to using an alternative specification of the instrument taking into account transition phases to the no school fees regime (see appendix). They are also robust to restricting the sample to individuals most affected by this educational reform. When we restrict our analysis to cohorts born before 1970

(8 years after the last state abolished secondary school fees), the results do not change qualitatively. One should note, however, that because of smaller sample sizes the instrument becomes rather weak.

4 Conclusion

In this paper we have estimated the causal effect of education on health behavior using the abolition of secondary school fees in Western Germany after World War II as an instrument. Different German states enacted this reform at different points in time generating a natural experiment. We found that the abolition of secondary school fees in intermediate and academic track had a small but statistically significant effect on educational attainment. While we find that better educated people smoke less, and have less weight problems, we do not find support for the notion that education *causes* better health outcomes.. Our paper therefore adds to a growing empirical literature questioning the presence of large positive effects of education on health.

References

- Albouy Valerie and Laurent Lequien (2008). Does compulsory schooling lower mortality? *Journal of Health Economic*. doi:10.1016/j.jhealeco.2008.09.003.
- Angrist, Joshua (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics* 19(1), 2-15.
- Arendt, Jacob Nielsen (2005). Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review* 24, 149-160.
- Clark, Damon and Heather Royer (2008). The effect of education on adult mortality and health: Evidence from the United Kingdom. Working paper.
- Currie, Janet and Enrico Moretti (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics* 118 (4), 1495-1531.
- Cutler, David M. and Adriana Lleras-Muney (2006). Education and health: Evaluating theories and evidence. NBER working paper 12352.
- De Walque, Damien (2007). Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics* 26, 877-895.
- Grimard, Franque and Daniel Parent (2007). Education and smoking: Were Vietnam draft avoiders also more likely to avoid smoking? *Journal of Health Economics* 26, 896-926.
- Grossman, Michael (2006). Education and non-market outcomes. In: *Handbook of the Economics of Education*, edited by: Eric Hanushek and Finis Welch, Elsevier Science.
- Imbens, Guido and Joshua D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica* 62(2), 467-475.
- Kenkel, Donald, Dean Lillard, and Alan Mathios (2006). The roles of high school completion and GED receipt in smoking and obesity. *Journal of Labor Economics* 24(3), 635-659.
- Oreopolous, Philip (2006). Estimating average and local average treatment effects of education when compulsory school laws really matter. *American Economic Review*, 96, 152-175.
- Pischke, Jörn-Steffen and Till von Wachter (2008). Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics* 90(3), 592-598.
- Riphahn, Regina (2008). The effect of secondary school fees on educational attainment. *LASER Discussion Papers - Paper No. 21*.
- Sander, William (1995). Schooling and Quitting Smoking. *Review of Economics and Statistics* 77(1), 191-199.
- Silles, Mary (2009). The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education Review* 28: 122-128.
- Staiger, Douglas and James H. Stock, 1997. Instrumental Variables Regression with Weak Instruments, *Econometrica* 65, 557-586.

Tables

Table 1: Sample description (means)

| | Microcensus 1999+2003 | |
|-----------------------------------|--------------------------------|--------------------------------|
| | Women | Men |
| Health Variables | | |
| Currently smoking | 0.28 (0.45) N=92,860 | 0.37 (0.48) N=91,694 |
| Ever smoked | 0.44 (0.50) N=92,015 | 0.60 (0.50) N=91,040 |
| Quitted smoking | 0.36 (0.48) N=40,497 | 0.37 (0.48) N=54,209 |
| Overweight | 0.35 (0.48) N=78,563 | 0.55 (0.50) N=79,691 |
| Obese | 0.10 (0.30) N=78,563 | 0.12 (0.33) N=79,691 |
| Individual Characteristics | | |
| Years of education | 10.00 (1.76) N=98,173 | 10.16 (1.86) N=97,818 |
| Year of birth | 1957.5 (12.73) N=109,315 | 1957.7 (12.64) N=108,351 |
| Fees | 0.23 (0.41) N=109,315 | 0.22 (0.41) N=108,351 |
| Transition | 0.17 (0.37) N=109,315 | 0.16 (0.36) N=108,351 |

Note: Std. deviation in parentheses.

Table 2: OLS results

| Coefficient on Years of Education | Currently Smoking | Ever Smoked? | Quitted Smoking? | Overweight | Obese |
|---|----------------------------------|----------------------------------|--------------------------------|----------------------------------|----------------------------------|
| Pooled | -0.040** (0.001) N=172,103 | -0.033** (0.001) N=172,732 | 0.034** (0.001) N=87,996 | -0.039** (0.001) N=147,871 | -0.018** (0.000) N=147,871 |
| Men | -0.043** (0.001) N=85,698 | -0.040** (0.001) N=85,102 | 0.029** (0.001) N=50,405 | -0.036** (0.001) N=74,638 | -0.018** (0.001) N=74,638 |
| Women | -0.036** (0.001) N=86,405 | -0.026** (0.001) 85,630 | 0.041** (0.002) N=37,591 | -0.044** (0.001) N=73,233 | -0.019** (0.001) N=73,233 |

Note: Regression also contain year of birth dummies, state of residence fixed effects, and gender dummies in the pooled sample. Standard errors in parentheses. ** 1% * 5% significance level.

Table 3: First stage IV results using dummy variable “Fees” as instrument. Effect of secondary school fees on years of schooling.

| Coefficient on Years of Education | Currently Smoking | Ever Smoked? | Quitted Smoking? | Overweight | Obese |
|-----------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Pooled | -0.103** (27.99) | -0.103** (28.24) | -0.102** (14.05) | -0.100** (23.55) | -0.100** (23.55) |
| Men | -0.111** (13.91) | -0.111** (13.75) | -0.074* (4.14) | -0.107** (11.67) | -0.107** (11.67) |
| Women | -0.096** (14.69) | 0.097** (15.11) | -0.149** (13.79) | -0.095** (12.61) | -0.095** (12.61) |

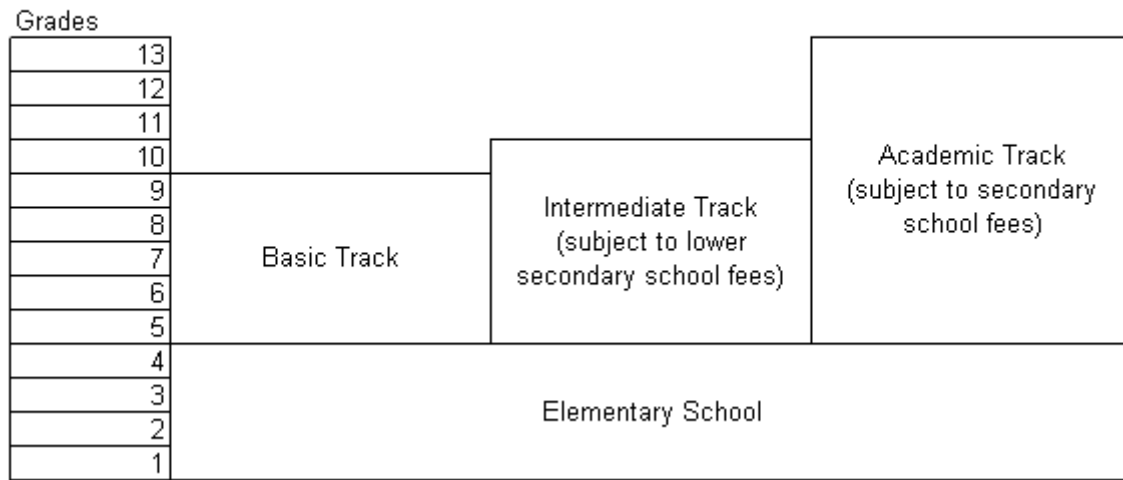
Note: In all columns the same equation using a dummy for never facing secondary school fees as instrument is estimated. Regression also contain year of birth dummies, state of residence fixed effects, and gender dummies in the pooled sample. First Stage F-Statistics on Secondary School Fees in Parentheses. ** 1% *5% significance level.

Table 4: Second stage IV results. Causal effect of education on health behavior and weight problems

| Coefficient on Years of Education | Currently Smoking | Ever Smoked? | Quitted Smoking? | Overweight | Obese |
|---|----------------------|------------------|---------------------|------------------|------------------|
| Pooled | -0.004 (0.048) | 0.023 (0.054) | 0.005 (0.076) | 0.117 (0.066) | 0.073 (0.045) |
| | N=172,103 | N=170,732 | N=87,996 | N=147,871 | N=147,871 |
| Men | 0.016 (0.068) | 0.016 (0.068) | -0.087 (0.144) | 0.057 (0.078) | 0.030 (0.059) |
| | N=85,698 | N=85,698 | N=50,405 | N=74,638 | N=74,638 |
| Women | -0.027 (0.069) | 0.052 (0.081) | 0.089 (0.085) | 0.188 (0.109) | 0.120 (0.070) |
| | N=86,405 | N=85,630 | N=37,591 | N=73,233 | N=73,233 |

Note: Regression also contain year of birth dummies, state of residence fixed effects, and gender dummies in the pooled sample. Standard errors in parentheses. ** 1% * 5% significance level.

Figure 1: The German School System



Source: Adapted from Riphahn (2008)

Appendix

Riphahn (2008) uses two different set of dummies to account for the abolition of school fees. In the paper, we use the dummy for being in a cohort experiencing no school fees during the whole educational career (column V in table 1). As an alternative specification we show results using a dummy for being in a transitional period. (column IV in table 1).

Table A1: First stage IV results for alternative definition of the instrument (“Transition.”). Effect of secondary school sees on years of schooling

| Coefficient on Years of Education | Currently Smoking | Ever Smoked? | Quitted Smoking? | Overweight | Obese |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| Pooled | -0.181*** (92.15) | -0.179*** (89.31) | -0.173*** (39.31) | -0.182*** (82.39) | -0.182*** (82.39) |
| Men | -0.167*** (32.69) | -0.165*** (31.59) | -0.152*** (17.51) | -0.170*** (30.48) | -0.170*** (30.48) |
| Women | -0.195*** (66.63) | -0.193*** (64.37) | -0.215*** (26.03) | -0.195*** (57.41) | -0.195*** (57.41) |

Note: Regression also contain year of birth dummies, state of residence fixed effects, and gender dummies in the pooled sample. First Stage F-Statistics on Secondary School Fees in Parentheses. ** 1% *5% significance level.

Table A2: Second stage IV results for alternative definition of the instrument “Transition”.
Causal effect of education on health behavior and weight problems

| Coefficient on Years of Education | Currently Smoking | Ever Smoked? | Quitted Smoking? | Overweight | Obese |
|---|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Pooled | 0.015 (0.027) N=172,103 | 0.011 (0.031) N=170,732 | -0.016 (0.047) N=87,996 | 0.019 (0.032) N=147,871 | 0.030 (0.024) N=147,871 |
| Men | 0.033 (0.045) N=85,698 | -0.012 (0.048) N=85,698 | -0.066 (0.070) N=50,405 | -0.023 (0.047) N=74,638 | 0.034 (0.038) N=74,638 |
| Women | 0.001 (0.032) N=86,405 | 0.027 (0.039) N=85,630 | 0.057 (0.063) N=37,591 | 0.058 (0.045) N=73,233 | 0.027 (0.030) N=73,233 |

Note: Regression also contain year of birth dummies, state of residence fixed effects, and gender dummies in the pooled sample. Standard errors in parentheses. ** 1% * 5% significance level.