Kemptner, Daniel; Jürges, Hendrik; Reinhold, Steffen

Conference Paper

Changes in Compulsory Schooling and the Causal Effect of Education on Health

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Economic Aspects of Health, No. B18-V1

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Kemptner, Daniel; Jürges, Hendrik; Reinhold, Steffen (2010) : Changes in Compulsory Schooling and the Causal Effect of Education on Health, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Economic Aspects of Health, No. B18-V1, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/37333

Changes in Compulsory Schooling and the Causal Effect of Education on Health

Daniel Kemptner[§], Hendrik Jürges[#], and Steffen Reinhold[*]

Preliminary Draft: December 16, 2009

Abstract

In this paper, we investigate the causal effect of years of schooling on health and health-related behavior in West Germany. We apply an instrumental variables approach using changes in compulsory schooling laws which took place from 1949 to 1969 as natural experiments. These law changed generate exogenous variation in years of schooling both across states and over time. We find evidence for a strong and significant causal effect of years of schooling on long-term illness for men but not for women. Moreover, we provide evidence of a causal effect of education on the likelihood of having weight problems for both sexes. On the other hand, we find little evidence for a causal effect of education on smoking behavior. Overall, our estimates suggest significant non-monetary returns to education with respect to health and health-related behavior.

Keywords: School reform, Health, Obesity, Smoking, Instrumental Variables

JEL-Codes: I12, I21

[§] DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany. Email: Daniel Kemptner <dkemptner@diw.de>
[#] MEA University of Mannheim, L13, 17, 68131 Mannheim, Germany. Email: juerges@mea.uni-mannheim.de
[*] MEA University of Mannheim, L13, 17, 68131 Mannheim, Germany. Email: reinhold@mea.uni-mannheim.de

1. Introduction

One of the most robust findings in both the economic and medical literature is the positive association between education and health (e.g. see the surveys by Cutler & Lleras-Muney 2008 or Grossman 2006). This relationship can be found in many countries, and it holds at different education levels and for various indicators of health. In fact, the relationship between education or socio-economic status in a wider sense and health is so ubiquitous that is often simply referred to as "the" gradient (Deaton 2003). However, the association between education and health behavior does not necessarily reflect a causal effect of education on health and there is now a lively debate, especially in the economics literature, whether this association (or how much of it) is causal.

This discussion is highly policy relevant because, among the components of socio-economic status (education, income, and occupation), education seems to be the one that – in the long run – appears to be most amenable to public policy interventions. Moreover, as noted by Deaton (2002), it is important to frame social policy in terms of health and wealth simultaneously. Improving one at the expense of the other involves a difficult and probably unnecessary trade-off. One possible policy instrument for improving health and wealth simultaneously is education. If education improves both wealth and health, giving people access to more and better education will – in the long run – be more successful than redistributing income or expanding public health care expenditures.

The question is thus: does education improve health? One *theoretical* explanation why more education might causally lead to better health is that education raises efficiency in health production (raises the marginal productivity of inputs), i.e. it increases an individual's productive efficiency (Grossman 1973). Just like more educated people are more productive in labor market activities they are likely to be more productive in non-market activities, which include the production of their own health (or the health of their children). Better educated people would thus found to be healthier even if inputs are held fixed.

Another (complementary) theoretical explanation is that education changes the inputs into health production (through information) and thereby increases allocative efficiency. Most prominently, education might change health behaviors such as smoking or bad nutrition habits (Rosenzweig & Schulz 1981, Cutler & Lleras-Muney 2007). It has been shown that education is associated with substantially reduced smoking and obesity rates, two important causes of premature deaths (Mokdad et al. 2004). For example, smoking rates in Germany are 48% among less educated men, as opposed to 20% among men with college degrees (Deutsches

Krebsforschungszentrum 2004). Education might improve health behaviors for several reasons. First, better educated people might be better informed about negative health consequences of smoking and overeating, either because they learned about these consequences in school, or because better educated people find it easier to obtain and evaluate such information (Kenkel 1991, Nayga 2000, de Walque 2007). Second, education could also influence health behavior through higher income, different social environments, a different sense of control, or an impact of education on time preferences, e.g. because schooling focuses students' attention on the future (Fuchs 1982, Becker & Mulligan 1997).

The aim of this paper is to estimate the causal effect of schooling on health exploiting variations in compulsory schooling laws in Western Germany. We follow the approach of Pischke and von Wachter (2005, 2008) who have used changes in years of compulsory schooling to estimate the causal effect of education on earnings. The authors found zero returns to schooling on wages. However, when estimating a reduced form equation regressing the body mass index on the change in compulsory schooling laws they found a significant effect. Our objective is to take this line of research further and assess the non-monetary returns of education by considering other health outcomes and also important dimensions of health-related behavior such as smoking. There is now a number of papers exploiting reforms in mandatory schooling laws across different countries, for instance the United States (Lleras-Muney 2005), the United Kingdom (Oreopolous 2006, Clark & Royer 2008, and Silles 2009), or France (Albouy & Lequien 2008). The results of these studies are mixed, showing negative effects of education on mortality (i.e. positive effects on health) in the US (Lleras-Muney 2005) but not in France (Albouy & Lequien 2008) or in the UK (Clark & Royer 2008). In contrast, strong positive effects on self-rated health in the UK are found by Oreopoulos (2006) and Silles (2009). Yet further studies for the UK, assessing the effect of compulsory schooling on morbidity using objective health measures such as blood pressure, BMI, or levels of inflammatory blood markers (Clark & Royer 2008, Jürges, Kruk & Reinhold 2009) find no effects.

Our study extends the existing literature by analyzing compulsory schooling laws in Germany to assess the causal effect of education in yet another economic environment where the *economic* returns to education at the margin of the educational reform were shown to be pretty low. In addition, our paper complements earlier work on the health effect of education in Western Germany, using the abolition of fees in academic track schools and academic track school constructions as different sources of exogenous variation in schooling (Jürges, Reinhold & Salm 2009, Reinhold & Jürges 2009).

The paper is structured as follows. In the next section we will describe the German school system and the changes in compulsory schooling laws as our institutional background. Section 3 explains which data we use and shows descriptive statistics. Section 4 describes our identification strategy and basic results. Here we also include results of additional analyses that support our basic results. Section 5 gives a brief summary and concludes the paper.

2. Institutional Background

In Germany, the federal states are responsible for educational policy. Still, the educational systems in all German states are very similar and thus share a lot of common features. Children generally enter primary school at the age of six. After (normally 4 years) in primary school, students attend one of three secondary school tracks which are in most cases taught at geographically separate schools. *Hauptschule* is the basic track leading to a basic school leaving certificate after grade 8 or 9. *Realschule* is a more demanding intermediate track which leads to a school leaving certificate after grade 10. Having finished school, both students from the basic track and the middle track usually start an apprenticeship or a school based vocational training. *Gymnasium* is the academic track leading to a general university-entrance diploma (*Abitur*) after grade 12 or 13. In addition to the three track types, some states also have comprehensive schools where all students are taught together (both types of school have some degree of internal tracking). Education in comprehensive school also leads to one of the three leaving certificates. However, this type of school is numerically unimportant for the cohorts studied in our paper. Some East German states have a middle school instead of basic track and intermediate track. Again, both types of leaving certificates are available at these schools. It is also noteworthy that, after being allocated to one of the three secondary school track, students rarely move between tracks, so that the track choice at the age of 10 usually has a strong implication for the entire life course (Dustmann 2004). The selection process itself depends on a mix of formal exams, grades in primary school, recommendations by the class teacher, and parental choice.

Since World War II, the German school system has undergone a number of reforms, among them the abolition of secondary school fees (Riphahn 2008), or the large-scale expansion of academic track schools (Jürges, Reinhold & Salm 2009). The reform we try to exploit in the present paper is the prolonging of compulsory basic track schooling from eight to nine years. Since education policy is in the hands of state governments, this reform has taken place at different points in time. Table 1 (taken from Pischke & von Wachter 2005)

contains, by federal state, the first year when all students were supposed to stay at least 9 years at school and the corresponding birth cohorts that were first affected by the reform. The first state to introduce a 9th grade in the basic track was Hamburg in 1949. The last state was Bavaria introducing a 9th grade in 1969. Interestingly, the data show a fairly consistent north-south pattern, with northern states being earlier reformers than southern states.

--- about here Table 1 (school reform year and cohorts) ---

3. Data and Descriptive Statistics

The data used in our study are taken from five years of the German Microcensus (1989, 1995, 1999, 2002, and 2003). The Microcensus is an annual, representative survey of one per cent of the households in Germany. Participation in the Microcensus is mandatory. However, the questions on health and health-related behavior are voluntary and are asked of a 45 percent random subsample only. The data contain, for each individual, sex and age, highest school degree and the current state of residence. In our analysis, we include cohorts born between 1930 and 1960 and currently living in the ten former West German states (excluding Berlin). We further restrict our analytical sample to individuals of German nationality and exclude individuals from the sample who have received their highest school degree in the former German Democratic Republic. Questions on health status (long-term illness and work-related disability), available only in the 2002 survey, were asked of working age individuals (16–65 years). We thus need to restrict our analysis to cohorts born between 1937 and 1960 when using the 2002 data on health status. We further exclude all individuals living in Hamburg from the 2002 survey since the change in compulsory schooling laws there first affected cohorts born in 1934.

Direct information on years in school as such is not available in the Microcensus. But the data provide the highest secondary school degree attained by each respondent. We use this information, together with both the number of years usually taken to obtain a certain degree and the compulsory schooling laws in the respective state, to impute years of schooling in each state, cohort and secondary school type. A dummy variable indicating whether the compulsory schooling laws require 8 or 9 years of schooling will serve as an instrument for an individual's years of schooling. Since geographical information is limited to state of residence at the time of the survey, we need to assume that a person has attended school in his state of residence. This raises concerns about the validity of the results of our analysis which we will address in a robustness check.

We measure health and health-related behavior of the respondents by a number of variables. Current health is measured by an indicator variable whether someone suffers from long-term illness (we do not have information on specific diagnoses) and whether this illness limits the type or amount of work they can do (work disability). As mentioned above, this information is only available in 2002 and only for working age individuals. Another measure of current health is the body mass index (BMI), derived from self-reported information on weight and height in the 1999 and 2003 surveys. We used BMI to construct two dummy variables indicating overweight (BMI greater than 25) and obesity (BMI greater than 30). Note that our definition of overweight includes obesity. Also note that BMI, overweight or obesity are can be interpreted as indicators of *future* health because they are significant risk-factors for future health problems, e.g. diabetes or cardiovascular disease. We measure health behavior by information on smoking. In 1989, 1995, 1999, and 2003, individuals were asked whether they were smoking currently and, if not, whether they had ever smoked. We use this information to identify quitters. Information on number of cigarettes smoked is only reported in rather broad bands. Analyses based on this measure delivers qualitatively similar results compared to our results using just binary information on smoking status. We generally pool all survey years which contain the same dependent variables, because each survey years represents an independent random sample.

--- about here Table 2 (data description) ---

Descriptive statistics are presented in Table 2. There are roughly 25,000 individuals per sex in the 2002 "current health" sample. These individuals were born between 1937 and 1960 and were on average 52 to 53 years old in 2002. Mean years in school were 9.6 years for men and 9.4 years for women. 20% of men and 15% of women were suffering from a long term illness. 17% of men and 13% of women suffered from a long-term illness that limited the amount or type of work they could do, i.e. they were work disabled. In the "BMI sample", interviewed in 1999 and 2003, there are about 50,000 observations per sex. Average BMI was 26.7 for men and 25.3 for women. 66% of men were overweight or obese compared to only 44% of women. Obesity rates were 16% for men and 13% for women. These gender differences can also be seen when looking at the kernel density estimates (Figure 1) showing that the distribution is shifted to the right for men. For both genders, one can see that a lot of the probability mass is concentrated in the BMI region around 25. For men, it even peaks at around those values.

---about her Figure 1---

For the analysis of smoking behavior, we have the largest samples with more than 120,000 observations for each sex. Overall, 63% of men and 39% of women reported having ever smoked. Current smoking rates are 36% for men and 24% for women. This implies that 44% of men (38% of women) in our sample who ever smoked have stopped smoking.

--- about here Table 3 (distribution of leaving certificates) ---

Table 3 shows the proportion of individuals in the birth cohorts 1930, 1940, 1950, and 1960 who graduated from each of the three secondary school tracks. These data clearly show the educational expansion in West Germany after the Second World War. More than three quarters of the population born in 1930 received at most basic track schooling. This proportion has diminished to less than 50% among those born in 1960. Moreover, the decline was much larger for women than for men. Both intermediate and academic track education has increased correspondingly. As a consequence, there is substantial variation across birth cohorts and states with respect to mean years of schooling (Table 3 for cohorts and Figure 2 for states).

---about here Figure 2---

Figure 2 shows a clear upward trends in mean schooling levels for all Western German states, from values below 9 years for the 1930 birth cohorts to values around or above 10 years for the 1960 birth cohorts. In addition to those time trends for the whole period, one can also see a jump of about 0.6 years in mean years of schooling at the time the states instituted the change in mandatory schooling laws. The pictures also illustrate the necessity to control for state-specific trends, as for instance Schleswig-Holstein seems to have a flatter incline in average educational attainment than Hamburg. In our basic specification, we will control for state-specific linear trends in addition to (common) time fixed effects to control for these different paths of the educational expansion across states.

4. Method and Results

We estimate both OLS and IV models. All equations are estimated for men and women separately and with robust standard errors which are clustered by state and year of birth. In the following subsection, we will first discuss the OLS results. Subsequently, we will present the IV results.

4.1 OLS results

Consider first the following model linking the health outcome to an individual's years of schooling:

$$H_i = \beta_0 + \beta_1 S_i + \beta_2 age^2 + (cohort * state)'\delta + \mu_{state} + \nu_{cohort} + \eta_{survey} + \varepsilon_i \qquad (1)$$

where H_i is a variable indicating some health outcome for individual i. S_i denotes individual i's years of schooling. μ_{state}, ν_{cohort} and η_{survey} are sets of fixed effects accounting for heterogeneity between federal states, birth cohorts and survey years. In addition, we add linear cohort-state interaction effects controlling for differential trends across states. Note that we cannot include a linear age effect because it is perfectly collinear with cohort and survey year combined. Yet we are able to include quadratic age term in order to address potentially non-linear age effect. When using the 2002 wave only, the survey-specific fixed effects and the quadratic in age are excluded from the right-hand side of equation (1).

Long-term illness and work disability. OLS regression results with respect to long term illness are shown in Table 4. The OLS coefficients indicate a significant association between current health and education in both sexes. It is, however, stronger for men than for women. One more year in school is linked with a reduction in the likelihood of suffering from a long term illness of 2.9 percentage points for men but only 1.2 percentage points for women. The association between years of schooling and long term illness is of substantive relevance considering that the overall prevalence rate of long term illness is 20% for men and 15% for women. The association between years in school and the probability of being work disabled is very pronounced and significant for both sexes, and it has about the same magnitude as the association of schooling and long-term illness.

--- about here Table 4 (OLS results) ---

BMI, overweight and obesity. The OLS estimates also indicate a significant association between years in school and body weight. Here, the association is much stronger for women than for men. One more year in school is linked with a decrease in BMI of 0.27 kg/m^2 for men and 0.47 kg/m^2 for women. In terms of overweight and obesity this means that each additional year in school is linked with a 3.1 (1.8) percentage point lower probability of being overweight (obese) for men, and a 4.7 (2.0) percentage point lower probability of being overweight (obese) for women. Again, these are fairly strong associations.

Smoking. Smoking behavior is also strongly related to schooling, but more so for men than for women. The OLS estimates show that one more year of schooling is linked with 2.6 percentage points lower rates of ever smokers among men but only 0.5 percentage points

lower rates of ever smokers among women. With respect to current smoking, the estimates suggest that one more year of schooling is associated with a decrease in the likelihood of being a smoker of 3.2 percentage points for men and of 2.1 percentage points for women. Further, the OLS estimates suggest a stronger association between years of schooling and the probability of quitting smoking for women than for men. One more year of schooling is associated with an increase in the probability of having quitted smoking of 4.1 percentage points among women and of 2.7 percentage points among men.

4.2 Instrumental variables results

The method of instrumental variables – using compulsory schooling laws discussed above as an instrument – allows dealing with the potential endogeneity of education and thus estimating a causal effect of education on health and health-related behavior. Our identification rests on the assumption that – conditional on covariates -- compulsory schooling laws or changes therein are uncorrelated with health outcomes and health behaviors, except through their effect via education. Technically, we assume that state education policies are independent of the error term of the second stage equation.

One particular concern is that state-specific trends in health or health behavior may be correlated with both the change in compulsory schooling laws and unobserved characteristics of the individuals that are themselves health determinants. If this was the case, the association between the instrument and the health outcome would not only be due to the association between the instrument and years of schooling. For example, there may be macroeconomic factors such as higher per capita growth which positively affect both the income of (parental) households and a state's resources which are available for reforms of the schooling system (such as financing one more school year for more than half of each cohort). If schools are financed by income taxes, there might also be a negative contemporaneous correlation between *net* parental household income and education expenses. In any case, as long as parental household income affects health (Case, Lubotsky & Paxson 2002, Reinhold & Jürges 2009), state-specific trends in economic development may compromise the validity condition. For this reason, we estimate the IV-model controlling for state-specific linear trends. These trends should also capture the effects of other measures of public policy which might be implemented together with the change in compulsory schooling laws. In additional robustness checks, we also investigate the robustness of our results to the inclusion of state-specific quadratic and cubic trends.

We estimate the following first stage equation:

$$S_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 age^2 + (cohort*state)'\delta + \rho_{state} + \tau_{cohort} + \lambda_{survey} + \nu_i, \quad (2)$$

where S_i denotes again years of schooling. Z_i is an instrumental variable indicating whether compulsory schooling laws require individual *i* to stay at least 8 or 9 years in school. When constructing the schooling variable S_i, we assigned 8 or 9 years of schooling to the basic track according to the applicable compulsory schooling laws. Thus, the association between the instrumental variable and the schooling variable (γ_1) arises by construction. ρ_{state}, τ_{cohort} and λ_{survey} are sets of fixed effects accounting for heterogeneity between federal states, birth cohorts and survey years. Inserting years of schooling predicted from equation (2), $\hat{S}_i$, into equation (1) yields the second stage equation:

$$H_i = \beta_0 + \beta_1 \hat{S}_i + \beta_2 age^2 + (cohort*state)'\delta + \mu_{state} + \nu_{cohort} + \eta_{survey} + \varepsilon_i, \quad (3)$$

When using the 2002 wave only, the survey-specific fixed effects and the quadratic in age are excluded from the right-hand side of equations (2) and (3)..

There are two concerns about the interpretation of our parameter of interest β_1 in equation (3). First, it is possible that the change in school attendance has not occurred sharply when a 9th grade of compulsory schooling was introduced (see Pischke and von Wachter 2005 for a discussion). As a consequence, years of schooling may be overstated for some students in the basic track shortly after the introduction of the 9th grade. This kind of one-sided measurement error would induce an attenuation bias even in IV-estimates so that our results have to be interpreted as lower bounds.

Second, to allows for heterogeneous effects of education on health, we need to assume monotonicity, i.e. we need to rule out that after the lengthening of mandatory schooling some students decide to go to school shorter than they would without the reform (no defiers). In our case it seems reasonable to assume that individuals with already high educational attainment (in intermediate or academic track) do not get less schooling because the mandatory years in basic track are raised by one year. With heterogeneous causal effects and when the monotonicity assumption holds, we identify the effect of education only on compliers, i.e. individuals whose number of years in school is affected positively by the reform (Angrist and Imbens, 1995). Moreover, since the number of years in school is a multi-valued treatment, the IV estimates are equal to a weighted average of all complier-specific causal effects.

--- about here Table 5 (IV results) ---

Table 5 shows the relevant first and second stage coefficients from our IV estimation. As indicated by the first stage results, mandatory schooling laws have high explanatory power with respect to years of schooling. The introduction of a compulsory 9th grade leads to an average increase of about 0.6 years in school. We use the Kleibergen-Paap weak identification test to assess the strength of our instruments (table 5). All test statistics are above 157 giving no rise to concerns about weak instrument problems.[1] It is also interesting to investigate for whom we identify the effect of education on health. Following Acemoglu and Angrist (2000) we estimate the average causal response weights, i.e. who are the compliers. Not surprisingly, we find that these weights are concentrated on individuals having nine years of schooling (96.5%) and much less weight on individuals with 10 years of schooling (2.3%) or higher levels of schooling. Thus, we identify an effect for individuals in the basic track getting nine instead eight years of schooling.

Long-term illness and work disability. IV estimates of the effect of schooling on health outcomes differ substantially between men and women. For men, we find a large and significant causal effect of years of schooling on health. Our results indicate that one more year of schooling reduces the likelihood of suffering from a long term illness by 4.1 percentage points for men (5% significance level). Furthermore, the estimates for men suggest that one more year of schooling leads to a reduction in the likelihood of work disability 3.2 percentage points, which is actually fairly close to the OLS result. In contrast to men, women appear not to gain from more schooling in terms of current health. The IV coefficients for women now have a *positive* sign but have become insignificant. We investigate the possibility that the substantial differences between the estimates for men and women arise because men with low levels of education tend to work more often in sectors where hard physical labor is performed, in particular in those older cohorts where labor force participation of women was much lower. For this reason, an additional year of schooling could be of particular relevance for their working conditions. Education is associated with non-manual labor occupations where dangers to physical health through exposure to injury, dangerous chemicals or adverse ergonomic conditions are reduced (WHO 2003). Furthermore, it could be the case that an individual with given health problems, as for instance back problems, would be able to work in a white-collar job but not in a blue-collar job anymore. If this was the case, then we would also expect more cases of work disability for blue-collar workers. To investigate this

[1] It has to be taken into account that the German Microcensus does not contain a direct measure of years of schooling. Since we assigned 8 or 9 years of schooling to the basic track according to the applicable compulsory schooling laws, the strong association between the instrumental variable and the schooling variable arises largely by construction.

hypothesis further, we constructed a new variable taking the value of 1 if the individual is working in a blue-collar job, 0 otherwise.[2] When conditioning on this additional variable in OLS regressions where long-term illness or work disability is the outcome we find that the coefficient on years of schooling is greatly reduced for men. Part of the education effect works through occupational choice. Furthermore, we find that education has also a positive causal effect on working in a white-collar job, albeit this estimate is only imprecisely estimated. Therefore, it seems that the gender differences in the effect of education on long term illness and work disability are mediated through the choice of occupational sector.

BMI, overweight, and obesity. For men's BMI, we find a significant effect of about the same size as the OLS results. One more year in school decreases male BMI by 0.3 kg/m^2. Among women, the effect approximately halves in size and becomes insignificant. Each year in school significantly reduces the probability of being overweight by 3.0 percentage points for men and 3.1 percentage points for women. Our effect sizes are consistent with the estimates for BMI in Pischke and von Wachter (table 3) where they find a *reduced* form coefficient on mandatory schooling laws of -0.160 kg/m^2 in the pooled sample of men and women. The relatively small effect on BMI and the strong effect on overweight can be explained by the fact that a lot the density mass is concentrated right around the critical threshold of a BMI of 25 (see Figure 1). The effect size is again in the range of the OLS coefficients for men but somewhat smaller for women. With respect to obesity, we obtain significant IV coefficients only for men but not for women.

Smoking behavior. In contrast to our health outcome measures and in contrast to the corresponding OLS results, our IV estimates do not suggest any significant effect of education on smoking behavior. This holds for ever smoking, current smoking, and quitting smoking and for both sexes. With one exception, the IV coefficients have become much smaller than the OLS estimates and some coefficients even change signs.

4.4 Robustness Checks

In the robustness checks, we address two concerns with our empirical strategy. First, we only have information on the current state of residence, but not on the state in which the individual has finished schooling. This leads to a mismeasurement of both the instrument and the endogenous regressor schooling. In one robustness check, we restrict our sample to four contiguous states changing the mandatory schooling laws at the same time. We argue that the

[2] These additional results are available from the authors upon request.

mismeasurement should be less severe in this restricted sample. Furthermore, we also probe our assumption that the trends in health in Western Germany can be captured by (common) cohort fixed effects and linear state-specific trends by estimating models adding also quadratic and cubic trends.

4.4.1 Restricting the sample to four contiguous states

A major concern about the validity of our results stems from the fact that geographical information is limited to state of residence at the time of the survey. Thus, we have to assume that a person has attended school in his current state of residence when constructing the instrumental variable as well as the schooling variable. This introduces a measurement error in both the instrument and the endogenous regressor, schooling. This is one source of attenuation bias discussed in Pischke and von Wachter (2005, 2008). We address this concern here by restricting our sample to four states for which this problem should be less severe.

--- about here Figure 3 (Map) ---

The four contiguous states Baden-Wuerttemberg, Rhineland-Palatinate, Hesse, and North Rhine-Westphalia introduced the 9th grade of compulsory schooling in 1967 (see Figure 3). In 1970, they represented 57% of the West German population. Since they do not border on a city state and contain large cities themselves, migration between the four states and other German states should be a minor issue. Furthermore, migration occurring *between* those four states does not matter since the 9th grade has been introduced *in the same year*. In a first step, we analyze the same birth cohorts as previously (1930-1960) but restrict our sample to the four contiguous states. When restricting the sample we obtain qualitatively very similar results compared to the whole sample.[3] Thus, we think that an attenuation bias due to migration between the states is not a major driver of our results.

Besides restricting our analysis to the four states, we now also restrict the sample to cohorts just affected and just not affected by the introduction of the 9th year of compulsory schooling in the basic track. In our original analysis, we consider the birth cohorts born between 1930 and 1960, and one might be concerned that it is difficult to adequately capture the time trends even in a flexible specification that we use. For this reason, we also restrict our sample to the last cohort for whom the change in compulsory schooling laws was not yet applicable and the first cohort who was supposed to attend 9 years of schooling in the basic track. The former cohort represents the control group and the latter cohort represents the treatment group. Our identification rests on the assumption that cohorts just before and just

after the cut-off point are very similar and differ only with respect to the treatment. In this case, we do not rely on estimating unobserved trends on the state level. In other words, we assume that the potential no-treatment outcomes of the two cohorts are identical. Technically, this means that we set cohort fixed effects to zero. We thus modify our IV estimation approach as follows: The first stage equation becomes:

$$S_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 age + \gamma_3 age^2 + \rho_{state} + \lambda_{survey} + \nu_i \quad (4)$$

The cohort-specific fixed-effects have been excluded from the right-hand side of equation (2). On the other hand, a linear age effect has been included since it is not captured anymore by cohort-specific and survey-specific fixed effects.

The second stage equation becomes:

$$H_i = \beta_0 + \beta_1 \hat{S}_i + \beta_2 age + \beta_3 age^2 + \mu_{state} + \eta_{survey} + \varepsilon_i \quad (5)$$

When we estimate the effect on health outcome that are only present in a single survey year of the German Microcensus, the survey-specific fixed effects, the linear age effect and the quadratic in age are excluded from the right-hand side of equations (4) and (5). The results are shown in Table 6. Because of the low number of clusters, we do not use clustered standard errors in the restricted sample.

--- about here Table 6 (robustness check) ---

The IV estimates for the restricted sample largely support our earlier findings – even though there are some differences. Naturally, the estimates are less precisely estimated when using the restricted sample. The estimated effect of education on long-term illness and work disability among men has increased substantially compared to the estimates for the unrestricted sample. For women, the point estimates have turned negative but they remain statistically insignificant. We also find substantially larger effects of years in school on BMI. Compared to the point estimates in the full sample the effects have more than doubled in size for men. For women they have more than tripled. Statistical significance is weak, however, which may be a results of the small size of the restricted sample. Similarly, despite substantially larger point estimates, we find statistically significant effect of years in school on the probability of being overweight or obese only for female obesity – which is in contrast to our results for the unrestricted sample.

Finally, considering smoking behavior, most effects remain insignificant although there are some changes with respect to the effect sizes. However, we find a significant effect

[3] Results available upon request.

of years of schooling on the probability of ever having smoked for women (at the 10% significance level) suggesting that among women, one more year of schooling causes a decrease in the probability of ever having smoked of 6.4 percentage points. Compared to the prevalence rate of 52.4% in the restricted sample, the estimated effect is relatively large.

4.4.2 More flexible state-specific trends

In this robustness check, we probe whether controlling for state-specific linear cohort trends in addition to (common) cohort fixed effects is sufficient to control for unobserved trends at the state level which are correlated with health and the decision to institute mandatory schooling laws. If the results do not change when one includes state specific quadratic or cubic trends one becomes more comfortable with the results. We present those additional results in table 7.

--- about here Table 7 (flexible state-specific trends)---

These results indicate that for men our results are very robust to the inclusion of more flexible state-specific trends. For long-term illness and work disability the coefficients maintain their significance levels and even slightly increase in size indicating a strong protective effect of education for men. For women, however, we still find basically a zero effect for long-term illness and work disability. These results are mirrored for BMI, and the probabilities of being overweight or obese. For men, the coefficients on education do not change much as one models the state-specific trends more flexible. However, for women, we now find a zero effect on weight problems. For smoking, we still do not find any significant results. These robustness checks suggest that there is no causal effect of education for women's health in our sample when using our instrument. However, our findings for men are strengthened.

5. Summary and Discussion

The present paper aims at contributing to the growing literature on identifying the *causal* link between education and health and health-related behavior. Economic theory has identified causal effects of education on health through several plausible channels: (a) education raises efficiency in health production; (b) education changes inputs into health production (through information) and thereby increases allocative efficiency; (c) education itself changes time preference (and thus inputs into health production) because schooling

focuses students' attention on the future; (d) education has an indirect effect mediated through higher income, occupational status, better housing, or healthier environmental conditions.

Numerous studies have indeed documented a strong positive empirical association between education and health. Interpretation of this correlation as causal is difficult, however, because education is most likely an endogenous variable, for instance because unobserved "pre-treatment" variables such as time preferences possibly drive both education and health behavior decisions, or because health (at younger ages, e.g. in early childhood) affects both educational achievement and later life health. Recent empirical work addresses causality issues head on using natural experiments such as exogenous changes in compulsory schooling laws for identification. Our paper is directly linked to this literature. Using data from several German Microcensuses, we exploit changes in years of compulsory schooling in West German federal states that took effect between 1949 and 1969 to estimate the causal effect of years in school on long-term illness, work disability, BMI (and overweight/obesity) and current and former smoking measured in 1989 to 2003.

We find evidence for a strong and significant negative causal effect of years of schooling on long-term illness and work disability among men. Our IV estimates are slightly larger than OLS but in contrast to some of the existing literature they do not become implausibly large. For women, however, we do not find any significant causal effects on health status. A possible explanation for this finding is that women with low levels of education are less likely to perform hard physical labor than men. Additional analyses show that the effect of education on long-term health and work disability for men seems to be largely mediated through its effect on the probability of being a blue-collar worker. We also find some evidence for negative causal effects of education on male body weight and somewhat less convincing support for a negative effect on female body weight.

We also address some possible concerns about the validity of our results. The main concern here is that, because we observe individuals only in their current state of residence and do not know the state where they actually finished school, migration between the states results in measurement error in both the instrument and schooling. Our robustness check – exploiting that fact that four large contiguous states have lengthened compulsory schooling in the same year -- suggests that migration is unlikely to compromise our estimation results. Additionally, we also address the concern whether unobserved state-specific trends which could be correlated with the instrument are biasing our results. In our most flexible specification, we allow for (common) cohort fixed effects and cubic state-specific cohort

trends. These robustness checks do not alter our main conclusions, but they put doubt on whether there is any effect of education on women's health in our sample.

Overall, our results are only partly in line with earlier studies analyzing the effect of compulsory school reforms on morbidity, e.g. the 1947 and 1973 UK education reforms. Oreopoulos (2006) or Silles (2009) find large significant positive effects of schooling on general self-rated health – which corresponds to our finding for long-term illness and work disability. In contrast to our results, Clark & Royer (2008) find no significant effect of education on BMI. However, they find a large and significant effect on the probability of being obese when considering the cohorts born from 1920-1950. Our findings on weight problems are in line with the study of Arendt (2005) who investigated changes in compulsory schooling laws in Denmark.

How can we interpret our results in the light of the theory? Based on our analysis, we can only partly distinguish between the four theoretical arguments for a causal effect of education on health mentioned above. A link via better health inputs in terms of less smoking is not supported by data. But, our findings on weight problems may shed some light on the mechanism of the causal relationship between education and health. It could well be that the more educated individuals are less likely to suffer from long-term illness or a work disability since they are less likely to be obese. Although we have not looked at income or wages in our paper, a link via higher income and thus favorable living conditions appears unlikely, because others – using an identical identification strategy – have found zero causal effects of mandatory schooling in Germany on wages (Pischke and von Wachter 2005, 2008).

Finally, it must also be noted that our parameters only identify the effect of education for compliers to the specific reforms of raising mandatory school leaving age. Interventions at other stages of the life-cycle or more specific interventions might have even stronger and more systematic effects on health outcomes and health behavior.

References

Acemoglu, D. and J. D. Angrist (2000): "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," In: National Bureau of Economics Macroeconomics Annual 2000. Editors: Ben S. Bernanke and Kenneth S. Rogoff. MIT Press Cambridge.

Albouy, V. and L. Lequien (2009): " Does compulsory education lower mortality?", Journal of Health Economics, Vol. 28, pp. 155-168.

Angrist, J. D. and G.W. Imbens (1995): "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." Journal of the American Statistical Association, Vol. 90 (130), pp. 431-442.

Becker, G.S. and C.B. Mulligan (1997): The endogenous determination of time preference, *Quarterly Journal of Economics* 112, 729-758.

Case, A., D. Lubotsky, and C. Paxson (2002): "Economic Status and Health in Childhood: The Origins of the Gradient". American Economic Review, Vol. 92, No. 5, pp. 1308-1334

Clark, D. and H. Royer (2008): "The Effect of Education on Adult Mortality and Health: Evidence from the United Kingdom", Working Paper.

Cutler, David, and Adriana Lleras-Muney (2008): Education and Health: Evaluating theories and evidence. In: The effects of social and economic policy on health. Editors: J. Newhouse, R. Schoeni, G. Kaplan, and H. Polack, forthcoming, Russell Sage Press.

Deaton , A. (2002): Policy implications of the gradient of health and wealth, *Health Affairs* 21 *(2)*, 13-30.

Deaton, A. (2003): Health, Income and Inequality. *NBER Reporter*, Spring 2003.

Deutsches Krebsforschungszentrum (2004): Rauchen und soziale Ungleichheit – Konsequenzen für die Tabakkontrollpolitik [Smoking and Health Inequality – Implications for tobacco control policy]. Heidelberg.

Dustmann, C. (2004): "Parental Background, Secondary School Track Choice, and Wages", Oxford Economic Papers, Vol. 56, pp. 209 - 230.

Fuchs, Victor (1982): Time Preference and Health: An Explanatory Study. pp. 93-120 in: VR Fuchs, Economic aspects of health, University of Chicago Press.

Jürges, H. E. Kruk, and S. Reinhold (2009): The effect of compulsory schooling on health – evidence from biomarkers. Working Paper. University of Mannheim.

Jürges, H, S. Reinhold, and M. Salm (2009): "Does schooling affect health behavior? Evidence from the educational expansion in Germany," IZA discussion paper 4330.

Kenkel, Donald (1991): Health behavior, health knowledge, and schooling. Journal of Political Economy, 99, 287-305.

Lleras-Muney, A. (2005): "The Relationship between Education and Adult Mortality in the United States", Review of Economic Studies, Vol. 72, pp. 189-221.

Mokdad, Ali H,, James S. Marks, Donna F. Stroup, and Julie F. Gerberding (2004): Actual causes of death in the United States, 2000. Journal of the American Medical Association, 291, 1238-1245.

Nayga, Rudolfo M. (2000): Schooling, health knowledge, and obesity. Applied Economics, 32, 815-832.

Oreopoulos, P. (2006): "Estimating average and local average treatment effects of education when compulsory schooling laws really matter," American Economic Review, Vol. 96 (1), pp. 152-175.

Pischke, J.-S. and T. von Wachter (2005): "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation", IZA Discussion Paper #1645.

Pischke, J.-S. and T. von Wachter (2008): "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation", The Review of Economics and Statistics, Vol. 90 (3), pp. 592-598.

Reinhold, S. and H. Jürges (2009): "Secondary school fees and the causal effect of schooling on health behavior." Health Economics, forthcoming.

Reinhold, S. and H. Jürges (2009b): "Parental Income and Child Health in Germany". MEA Discussion Paper 175-2009.

Riphahn, Regina (2008): The effect of secondary school fees on educational attainment. LASER Discussion Papers - Paper No. 21.

Rosenzweig, MR and Schulz. TP (1981): Education and Household Production of Child Health. In: Proceedings of the American Statistical Association (Social Statistics Section), 382-387.

Silles, M. (2009): "The Causal Effect of Education on Health: Evidence from the United Kingdom", Economics of Education Review 28 (1), 122-128.

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.

Table 1 – Introduction of a 9th grade in the basic track of secondary school

	First year in which all students were supposed to graduate no sooner than after nine years of compulsory schooling	First birth cohort affected by the change in compulsory schooling laws
Hamburg	1949	1934
Schleswig-Holstein	1956	1941
Bremen	1958	1943
Lower Saxony	1962	1947
Saarland	1964	1949
North Rhine-Westphalia	1967	1953
Hesse	1967	1953
Rhineland-Palatinate	1967	1953
Baden-Wuerttemberg	1967	1953
Bavaria	1969	1955

Source: Pischke and von Wachter, 2005

Table 2: Sample means (standard deviation in parentheses, number of observations below)

Variable	Sex	Microcensus 2002 Cohorts born between 1937 and 1960	Microcensus 1999, 2003 Cohorts born between 1930 and 1960	Microcensus 1989, 1995, 1999, 2003 Cohorts born between 1930 and 1960
Age	Men	52.48	54.76	50.13
		(7.10)	(9.10)	(10.27)
		24,099	49,843	122,178
	Women	52.63	55.19	50.57
		(7.11)	(9.17)	(10.31)
		24,624	50,560	124,674
Years in school	Men	9.63	9.50	9.42
		(1.82)	(1.84)	(1.80)
		23,128	48,640	119,461
	Women	9.39	9.20	9.17
		(1.62)	(1.59)	(1.55)
		23,479	49,225	121,541
Long-term illness	Men	0.20	-	-
		(0.40)		
		24,099		
	Women	0.15	-	-
		(0.36)		
		24,624		
Work disability	Men	0.17	-	-
		(0.37)		
		23,791		
	Women	0.13	-	-
		(0.33)		
		23,369		
BMI	Men	-	26.70	-
			(3.66)	
			49,843	
	Women	-	25.25	-
			(4.36)	
			50,560	
Overweight	Men	-	0.66	-
			(0.47)	
			49,843	
	Women	-	0.46	-
			(0.50)	
			50,560	
Obesity	Men	-	0.16	-
			(0.36)	
			49,843	
	Women	-	0.13	-
			(0.34)	
			50,560	
Ever smoked	Men	-	-	0.63
				(0.48)
				122,178
	Women	-	-	0.39
				(0.49)
				124,674
Currently smoking	Men	-	-	0.36
				(0.48)
				122,178
	Women	-	-	0.24
				(0.43)
				124,674
Quitted smoking	Men	-	-	0.44
				(0.50)
				77,566
	Women	-	-	0.38
				(0.49)
				48,501

Table 3: Distribution of leaving certificates in percent, by birth cohorts

		1930	1940	1950	1960
Basic Track	Men	76.8	69.1	58.8	47.7
	Women	82.1	73.0	63.0	40.7
Intermediate	Men	12.3	14.6	17.2	22.7
	Women	12.9	19.8	22.4	35.5
Academic	Men	10.9	16.3	23.9	29.6
	Women	4.9	7.2	14.6	23.8

Table 4 –OLS results

Dependent variable		Sex	Coefficient on years in school	Standard Error	Number of observations
Current Health	Long term illness	Men	-0.029***	0.001	23,128
		Women	-0.012***	0.002	23,479
	Work disability	Men	-0.029***	0.001	22,858
		Women	-0.013***	0.001	23,266
Weight	BMI	Men	-0.271***	0.010	48,640
		Women	-0.465***	0.013	49,225
	Overweight	Men	-0.031***	0.001	48,640
		Women	-0.047***	0.002	49,225
	Obesity	Men	-0.018***	0.001	48,640
		Women	-0.020***	0.001	49,225
Smoking	Ever smoked	Men	-0.026***	0.001	119,461
		Women	-0.005**	0.002	121,541
	Currently smoking	Men	-0.032***	0.001	121,318
		Women	-0.021***	0.002	124,314
	Quitted smoking	Men	0.027***	0.001	76,037
		Women	0.041***	0.002	47,551

Note: Standard errors clustered on cohort*state level. * p<10 percent, ** p<5 percent, *** p<1 percent. All regressions include fixed effects for year of birth, state of residence, and an interaction of state of residence and a linear cohort trend. If several survey years have been pooled, fixed effects for survey year and the quadratic in age are included. . Regressions with weight outcomes include height as additional control variable..

Table 5: Results of 2SLS regressions

Dependent variable		Sex	First Stage Coefficient on instrument	Second stage coefficient on years in school	Number of observations
Current Health	Long term illness	Men	0.656***	-0.041**	23,128
			(0.054)	(0.017)	
			160.3		
		Women	0.579***	0.010	23,479
			(0.047)	(0.017)	
			161.8		
	Work disability	Men	0.655***	-0.032**	22,858
			(0.055)	(0.015)	
			157.7		
		Women	0.580***	0.021	23,266
			(0.047)	(0.016)	
			160.5		
Weight	BMI	Men	0.595***	-0.301**	48,640
			(0.040)	(0.121)	
			238.0		
		Women	0.663***	-0.194	49,225
			(0.033)	(0.133)	
			380.2		
	Overweight	Men	0.595***	-0.030**	48,640
			(0.040)	(0.015)	
			238.0		
		Women	0.663***	-0.031**	49,225
			(0.033)	(0.015)	
			380.2		
	Obesity	Men	0.595***	-0.030**	48,640
			(0.040)	(0.014)	
			238.0		
		Women	0.663***	-0.004	49,225
			(0.033)	(0.010)	
			380.2		
Smoking	Ever smoked	Men	0.615***	-0.011	119,461
			(0.025)	(0.011)	
			620.8		
		Women	0.655***	0.010	121,541
			(0.021)	(0.010)	
			952.3		
	Currently smoking	Men	0.616***	-0.005	121,318
			(0.025)	(0.010)	
			631.5		
		Women	0.654***	-0.000	124,314
			(0.021)	(0.009)	
			975.1		
	Quitted smoking	Men	0.668***	-0.005	76,037
			(0.030)	(0.011)	
			536.2		
		Women	0.690***	0.015	47,551
			(0.031)	(0.012)	
			507.8		

Note Clustered standard errors (on cohort*state level) in parentheses. In addition, we report Kleibergen-Paap rK Wald F statistics in italics. All regressions include fixed effects for year of birth, state of residence, and an interaction of state of residence and a linear cohort trend. If several survey years have been pooled, fixed effects for survey year and the quadratic in age are included. . Regressions with weight outcomes include height as additional control variable. * p<10 percent, ** p<5 percent, *** p<1 percent.

Table 6: Robustness check (sample restricted to Baden-Wuerttemberg, Rhinland-Palatinate, Hesse and Northrhine-Westphalia)

Dependent variable		Sex	OLS	2SLS First stage	2SLS Second Stage	N obs.
Current Health	Long-term illness	Men	-0.028***	0.589***	-0.081**	1,097
			(0.006)	(0.111)	(0.037)	
		Women	-0.011*	0.716***	-0.013	1,199
			(0.006)	(0.093)	(0.028)	
	Work disability	Men	-0.027***	0.592***	-0.074**	1,095
			(0.005)	(0.111)	(0.034)	
		Women	-0.010*	0.721***	-0.014	1,194
			(0.005)	(0.093)	(0.025)	
Weight	BMI	Men	-0.118**	0.512***	-0.721	1,874
			(0.050)	(0.125)	(0.526)	
		Women	-0.412***	0.692***	-0.751*	1,772
			(0.059)	(0.114)	(0.438)	
	Overweight	Men	-0.010*	0.512***	-0.061	1,874
			(0.006)	(0.125)	(0.067)	
		Women	-0.041***	0.692***	-0.064	1,772
			(0.007)	(0.114)	(0.048)	
	Obesity	Men	-0.013***	0.512***	-0.020	1,874
			(0.004)	(0.125)	(0.049)	
		Women	-0.015***	0.692***	-0.072**	1,772
			(0.004)	(0.114)	(0.032)	
Smoking	Ever smoked	Men	-0.033***	0.507***	0.021	4,525
			(0.004)	(0.080)	(0.041)	
		Women	-0.018***	0.647***	-0.064*	4,447
			(0.005)	(0.069)	(0.034)	
	Currently smoking	Men	-0.047***	0.512***	0.023	4,585
			(0.004)	(0.079)	(0.042)	
		Women	-0.036***	0.650***	-0.037	4,513
			(0.004)	(0.069)	(0.031)	
	Quitted smoking	Men	0.041***	0.640***	-0.018	3,029
			(0.005)	(0.092)	(0.040)	
		Women	0.047***	0.677***	-0.009	2,346
			(0.006)	(0.093)	(0.042)	

Robust standard errors in parentheses. All regressions include fixed effects for state of residence. If several survey years have been pooled, fixed effects for survey year and the quadratic in age are included. Regressions with weight outcomes include height as additional control variable. * p<10 percent, ** p<5 percent, *** p<1 percent

Table 7: Robustness check (linear, quadratic, and cubic state-specific trends)

Dependent variable		Sex	+linear state trends	+quadratic trends	+cubic trends	N obs.
Current Health	Long-term illness	Men	-0.041**	-0.039**	-0.045**	23,128
			(0.017)	(0.019)	(0.018)	
		Women	0.010	0.013	0.024	23,479
			(0.017)	(0.024)	(0.027)	
	Work disability	Men	-0.032**	-0.037**	-0.045***	22,858
			(0.015)	(0.017)	(0.017)	
		Women	0.021	0.036	0.040	23,266
			(0.016)	(0.022)	(0.025)	
Weight	BMI	Men	-0.301**	-0.306**	-0.356**	48,640
			(0.121)	(0.136)	(0.146)	
		Women	-0.194	-0.050	0.149	49,225
			(0.133)	(0.150)	(0.168)	
	Overweight	Men	-0.030**	-0.034**	-0.039**	48,640
			(0.015)	(0.015)	(0.016)	
		Women	-0.031**	-0.016	-0.005	49,225
			(0.015)	(0.016)	(0.020)	
	Obesity	Men	-0.030**	-0.028*	-0.021	48,640
			(0.014)	(0.015)	(0.016)	
		Women	-0.004	0.007	0.023*	49,225
			(0.010)	(0.012)	(0.013)	
Smoking	Ever smoked	Men	-0.011	-0.012	-0.004	119,461
			(0.011)	(0.013)	(0.014)	
		Women	0.010	0.014	0.015	121,541
			(0.010)	(0.011)	(0.011)	
	Currently smoking	Men	-0.005	0.005	-0.001	121,318
			(0.010)	(0.012)	(0.014)	
		Women	-0.000	0.001	0.004	124,314
			(0.009)	(0.010)	(0.012)	
	Quitted smoking	Men	-0.005	-0.019	-0.006	76,037
			(0.011)	(0.013)	(0.014)	
		Women	0.015	0.015	0.011	47,551
			(0.012)	(0.015)	(0.017)	

Robust standard errors in parentheses. All regressions include fixed effects for state of residence. If several survey years have been pooled, fixed effects for survey year and the quadratic in age are included. Regressions with weight outcomes include height as additional control variable. * p<10 percent, ** p<5 percent, *** p<1 percent

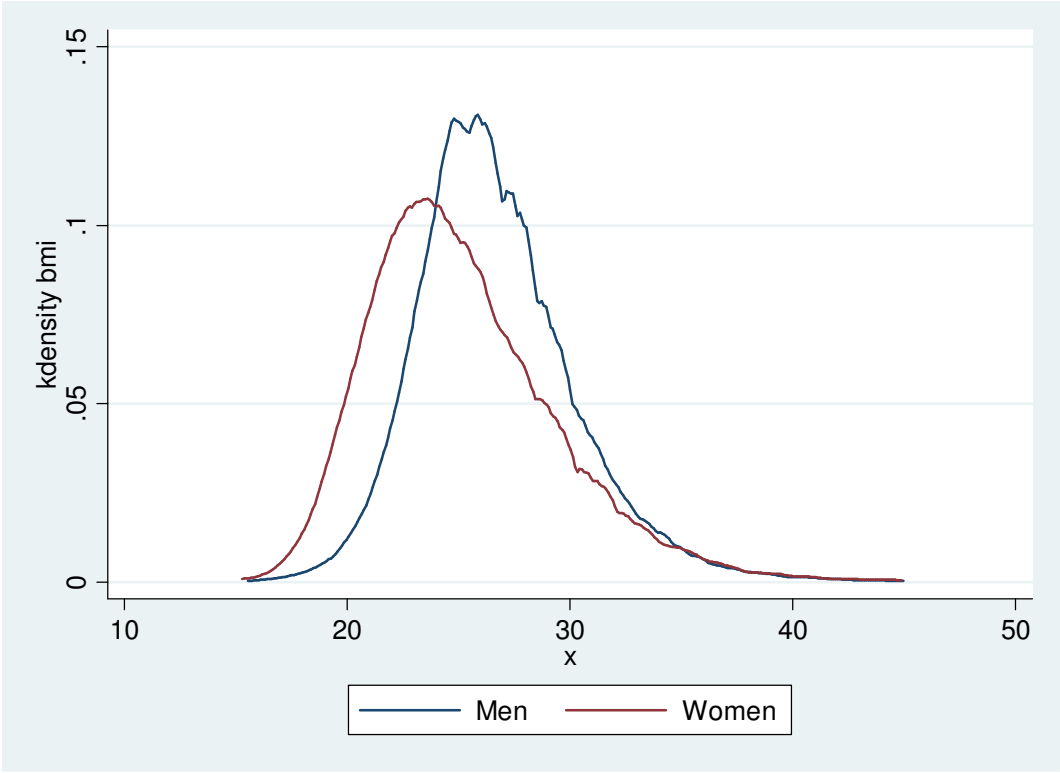

Figure 1: Kernel density estimate of BMI.

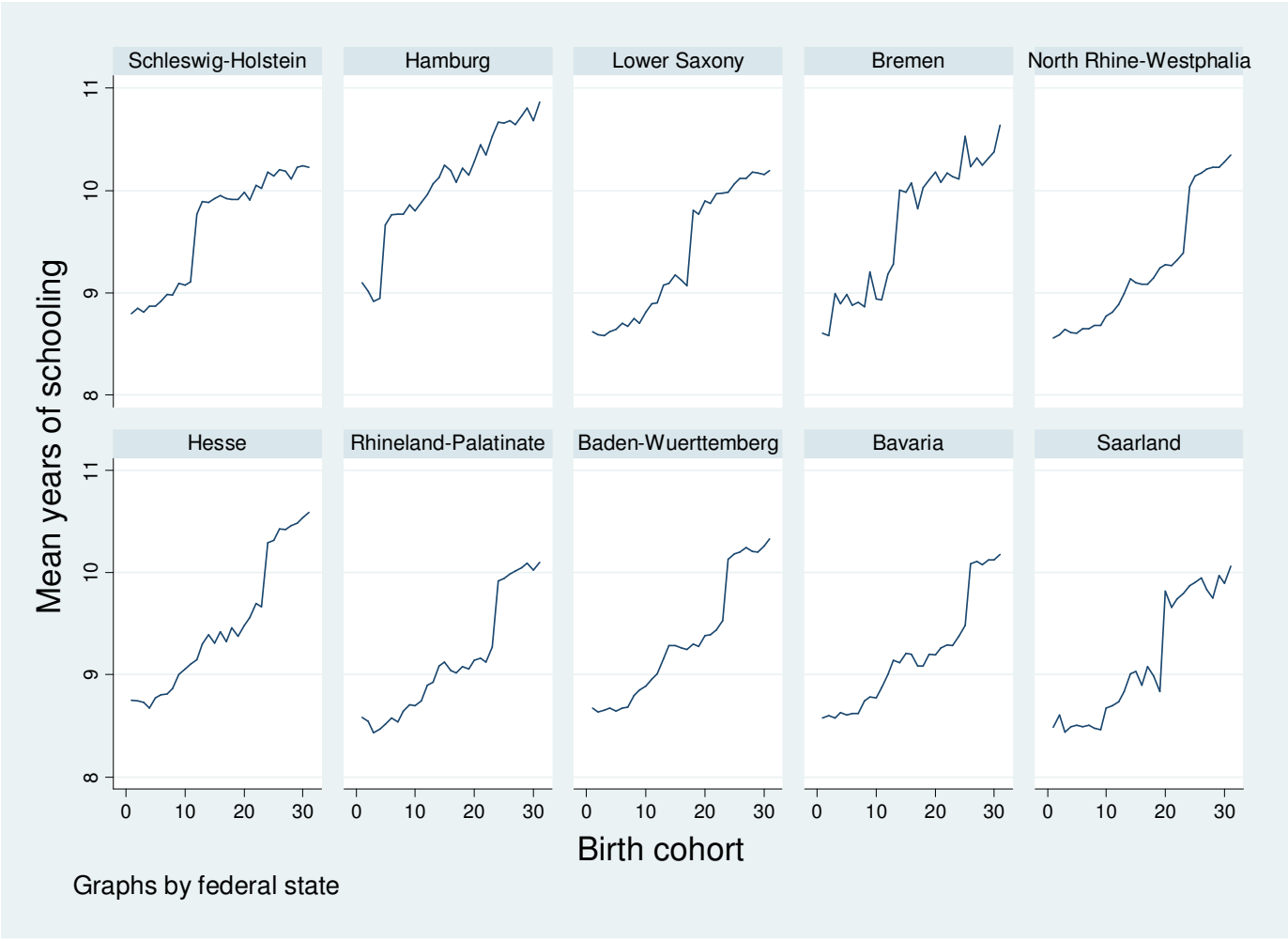

Figure 2: Mean number of years of schooling by birth cohort and federal state

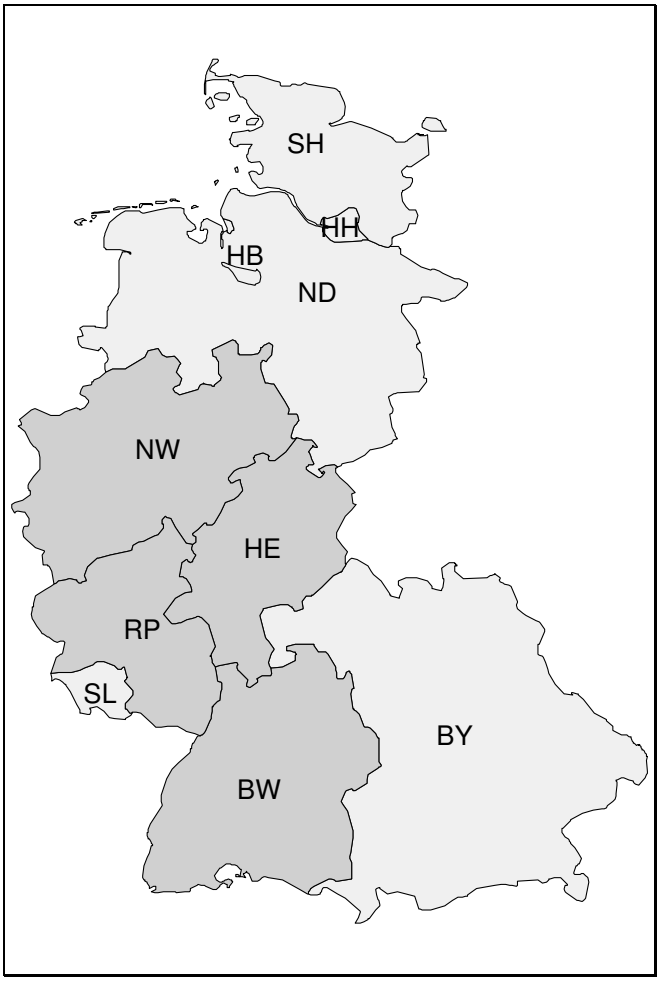

Figure 3: Map of West Germany showing the four contiguous states that lengthened compulsory schooling in the same year (NW=Northrhine-Westphalia, HE=Hesse, RP=Rhineland-Palatinate, BW=Baden-Wuerttemberg)