Is additional schooling worthless? Revising the zero returns to compulsory schooling in Germany

Kamila Cygan-Rehm

This version: April 18, 2018

Abstract: This study estimates the effect of compulsory schooling on earnings. For identification, I exploit a German reform that extended the duration of secondary schooling in the 1960s. I find that hourly wages increase by 6%-8% per additional year of schooling. This result challenges a study by Pischke and von Wachter (2008) who find zero returns to schooling using the same survey data and reform. I show that their estimates suffer from unconsidered institutional details. A complementary analysis using social security records confirms significant effects on earnings, but yields no effects on employment and eligibility for public transfers.

JEL Classification: I21, I26, J31

Contact: Kamila Cygan-Rehm, Dept. of Economics, Friedrich-Alexander University Erlangen-Nürnberg, Lange Gasse 20, 90403 Nuremberg, Germany, Email: kamila.cygan-rehm@fau.de.

The author would like to thank Deborah Cobb-Clark, Christina Gathmann, Mathias Hübener, Daniel Kühnle, Steve Pischke, Regina T. Riphahn, Hendrik Schmitz, and participants of the annual LASER meeting 2017 for their comments and advice on this paper. Heiko Stüber provided help with the SIAB data and Lena Höcker excellent research assistance. The author acknowledges financial support by Joachim Herz Stiftung (grant number 600047).

1 Introduction

A rapid increase in educational attainment was one of the most remarkable developments during the twentieth century. This educational expansion presumably had various reasons such as advances in technology, new requirements of the global economy, and demographic trends. Most developed countries addressed the new challenges by extending compulsory schooling requirements assuming that education plays a central role in modern labor markets. Since the spread of the seminal concept treating education as a capital investment (Schultz, 1961; Becker, 1962; Mincer, 1974), extensive empirical research has investigated the value of education, to determine whether governments and individuals invest optimally.

However, obtaining causal estimates of returns to schooling is difficult because various sources of bias might lead to both overestimated and underestimated effects from ordinary least-squares (OLS) regressions (Card, 1999). To approach a major challenge of omitted variable bias, economists have exploited various sources of (arguably) exogenous variation in education such as compulsory school attendance laws, schooling reforms, and accessibility of educational institutions with key references being Angrist and Krueger (1991), Harmon and Walker (1995), and Card (1995), respectively. The landmark studies found that an additional year of schooling increases wages by 10%-16 %. However, subsequent research often questioned such large returns by providing smaller estimates and occasionally even zero effects from compulsory schooling reforms (see e.g., Meghir and Palme, 2005; Pischke and von Wachter, 2008; Grenet, 2013).

The literature includes also several examples of prominent results that have been later revised. For instance, exploiting increases of minimum school leaving age (MSLA) in Britain, Oreopoulos (2006) estimated wage returns of 15%. Using his same sample and specification, Devereux and Hart (2010) could not replicate this result and arrived at returns of 7% instead. Additional adjustment for gender further rebutted the estimates to insignificant 2%-3%. In the latest installment in that debate, Dolton and Sandi (2017) provide robust estimates of 6% for men. For the U.S., the study by Acemoglu and Angrist (2000), who document wage effects of 6%-11% using variation in child labor laws and compulsory attendance laws over time and across states, has been recently challenged by Stephens and Yang (2014). The authors show that the prior estimates are entirely driven by differential developments across states (e.g., regarding school quality improvements). Thus, the existence of important economic returns to schooling remains one of the much disputed and controversial topics within economics.

This paper adds to this literature by revising estimates of the wage returns to schooling in the biggest European economy - Germany. I largely build on a previous study by Pischke and von Wachter (2008) who exploit the staggered extension of compulsory schooling duration from eight to nine years ($C9$) across West German states after World War II. Using survey data from the Qualification and Career Survey (QaC) and the Micro Census (MC), the authors conclude about zero wage returns to compulsory schooling in Germany[1]; a finding that has since been widely accepted and cited. Indeed, their main two-stage least-squares (2SLS) estimate based on the QaC data is statistically insignificant. However, the effect's magnitude implies an economically not trivial wage return of 5.8% per schooling year.[2] Moreover, earlier results in Pischke and von Wachter (2005) imply that the estimate nearly doubles and becomes significant after

[1] Kamhöfer and Schmitz (2016) arrive at the same conclusion and find likewise an insignificant effect on cognitive skills using a smaller sample from the German Socio-economic Panel (SOEP).

[2] The coefficient is 0.058 with a standard error of 0.038. Note that the 2SLS estimate does not exactly match the reported reduced-form and first-stage coefficients of 0.010 and 0.190, respectively as their ratio is 0.053.

controlling for squared instead of linear terms for state-specific cohort trends. In support of their conclusion about zero returns, Pischke and von Wachter (2008) show complementary estimates from the MC, which are smaller in magnitude and preciser, though the confidence intervals still include moderate positive returns. The authors mention also that they find a statistically significant return of 7.4%[3] when they restrict the MC sample to men, but attribute this finding to sample variability. The sensitivity of the previous estimates makes the conclusion that compulsory schooling in Germany yields no economic benefits rather controversial and encourages a reassessment.[4]

The main contribution of this paper is to provide more robust and precise estimates of returns to compulsory schooling in Germany by addressing some of the challenges that complicated the inference in Pischke and von Wachter (2008). I do so by implementing relatively minor, but substantively important, adjustments to their sample and specification. The are three major modifications: First, while Pischke and von Wachter (2008) study individuals born between 1930-1960, I exclude those born before 1945 because they were affected by wartime distortions and temporary extensions of compulsory schooling preceding the actual passage of the $C9$ laws. Second, I also exclude the "pivotal" birth cohorts that were exactly at the threshold to encounter the $C9$ reform due to the cutoff regulations for school enrollment. Third, I account for the parallel implementation of short school years (SSY) in several states as they also affected schooling duration (for details see Pischke, 2007). I demonstrate that if ignored, these institutional specifics confound the estimates and explain the insignificant estimates from

[3] The two-sample 2SLS point estimate is 0.074 with a standard error of 0.038.

[4] The lack of wage returns is also surprising in light of related research that establishes significant non-pecuniary effects of the German reform in terms of health (Kemptner et al., 2011), fertility (Cygan-Rehm and Maeder, 2013), and the intergenerational transmission of education (Piopiunik, 2014).

the original study.

Re-analyzing the QaC data, I find wage returns of 6%-8%. The sizable and statistically significant effects remain robust in various specification tests. I devote particular attention to controlling for the underlying trends in schooling and earnings. An auxiliary analysis on high-quality data from social security records confirms that the $C9$ reform significantly affected earnings, but had no substantial impact on employability and related outcomes such as eligibility for public transfers.

The paper proceeds as follows. Section 2 gives relevant institutional details. Section 3 introduces my empirical strategy and Section 4 describes the data. Section 5 shows the results. Section 6 explores the empirical robustness of my main findings and Section 7 concludes.

2 Instututional background

Traditionally and predominantly until today, the secondary schooling in Germany has a tripartite structure that distinguishes between basic (*Hauptschule*), intermediate (*Realschule*), and high schools (*Gymnasium*).[5] Students receive a referral to one of three tracks after four grades of primary schooling, typically at the age of ten. The tracking depends on various criteria, which differ by state. Usually, the primary school teachers make a recommendation, but in several states, parents do not have to comply. The tracks differ substantially in the academic content of the curriculum, thereby preparing for different professional careers. The basic track lasts until grade 8 or 9 and prepares for an apprenticeship. The intermediate track comprises 10 grades

[5] There are also alternative school types such as comprehensive schools without tracking (*Gesamtschule*) or schools for children with special needs (*Sonderschule, Förderschule*), but the vast majority of cohorts under this study still participated in the traditional tripartite system.

and qualifies for an apprenticeship or training in white collar jobs. A high school certificate after 12 or 13 grades gives access to academic education in colleges or universities. During the 1950s, about three-quarters of 11-year-olds attended the basic, about 10% the intermediate, and 15% the academic track. However, due to educational expansion, the fraction of basic track dropped to about 50% within two decades (STBA, 2006).

While the authority for educational policies lies with the federal states, all states negotiate and set framework agreements to ensure the comparability of school systems. The duration of compulsory schooling was one of the much disputed and controversial topics in the German educational debate after World War II (e.g., Leschinsky and Roeder, 1980). Figure 1 illustrates how the implementation of the ninth compulsory school year developed over time and across states. The Nazi regime centralized the education system and required at least eight years of schooling. Shortly after the war, Hamburg (1946) and Schleswig-Holstein (1947) re-introduced a compulsory ninth grade ($C9$), mandated there already before the war. The remaining states initially stuck to an eight-year standard, but in some states (e.g., Bremen and North Rhine-Westphalia) students in the basic track could voluntary attend the ninth or even tenth grade in selected schools.

In the postwar period, the weak labor market and the shortage of apprenticeship positions for school leavers became leading arguments for a nationwide extension of compulsory schooling (Petzold, 1981). Indeed, during the early 1950s, several states issued laws that allowed the authorities (at the state or local level) to prolong school attendance. The exact content of these laws differed by state. Some of the extensions were limited to specific birth cohorts, predominantly those first enrolled during the war. Several of the laws explicitly conditioned on insufficient number of apprenticeships relative to school leavers. However, due to the recovering labor market in the mid 1950s, the demand for apprentices increased, which virtually suspended the conditional schooling extensions. Only two further states - Saarland (1958) and Bremen (1959) - decided to permanently introduce the $C9$ year still in the 1950s. By then, the political discussion already shifted to educational arguments such as improving students' physical and psychological readiness for the labor market, and the quality and maturity of their occupational choices.

Finally, in 1964, the prime ministers of all states agreed that compulsory schooling should last nine years nationwide (*Hamburg Accord*). North Rhine-Westphalia, Hesse, Rhineland-Palatinate, and Baden-Wurttemberg eventually implemented the $C9$ reform in 1967, simultaneously with a shift of the starting date of the school year from spring to fall. The change in schedule was completed within two short school years (SSY) that actually jointly lasted only 16 months (for details see Pischke, 2007). Save for Hamburg, Lower Saxony, and Bavaria all school children at that time experienced compressed school years. Although the SSY shortened the instructional time in the attended grade, they did not affect the curriculum. Bavaria postponed the implementation of the $C9$ grade until 1969. Generally, the timing of the reform varied not only across states but also more locally, mainly because of challenges related to additional demand for teachers and classrooms (Leschinsky and Roeder, 1980). For example, Hesse enacted the $C9$ year step-wise starting in 1964 with counties that already fulfilled the organizational requirements.

Note that my description of the $C9$ reform largely follows Leschinsky and Roeder (1980) who provide different implementation dates for Schleswig-Holstein, Hamburg, Bremen, and Saarland compared to Pischke and von Wachter (2008). Given that I found striking inconsistencies regarding the timing of the reform across various sources, I reviewed the original state laws and official statistics on the actual ninth grade attendance from STBA (2006). Table 1 illustrates the incidence of ninth grade attendance in the basic track between 1954 and 1966.[6] For example, looking at Saarland (last column), we observe that the reform became effective in 1958 as the ninth grade attendance jumps to 99% from 0% in previous year, which exactly matches the timing from Leschinsky and Roeder (1980) and the original legislation. The numbers for Schleswig-Holstein (first column) also contradict the date from Pischke and von Wachter (2008), though the implementation in this state occurred gradually over several years. The revised timing of the reform should not, however, be decisive for the analysis as it affects relatively small states.

Table 1 uncovers also other interesting patterns. First, even states that introduced the reform relatively early such as Hamburg and Schleswig-Holstein, reached a ninth grade attendance of around 80%. Second, the temporary laws of the early 1950s, which allowed local schooling extensions depending on the labor market situation, remained at the longest effective in Rhineland-Palatinate (until 1957). The data point to a notable phenomenon in this state. In Lower Saxony and North Rhine-Westphalia, these extensions became irrelevant in 1954, and were hardly ever implemented in Bavaria. In several states, we also observe ninth graders under the eight-year standard, which are mostly voluntary extensions and reporting issues[7].

Generally, the reform directly affected the duration of schooling among students who would

[6] Similar statistics for earlier and later years are not available (STBA, 2006). These ratios should be interpreted with caution as they show the number of students in their ninth schooling year compared to students in their eight year, but in previous calendar year. Thus, Table 1 includes regional mobility and switches between tracks.

[7] For example, until 1960, these numbers include students from a special type of classes that actually prepared for the intermediate degree, but were integrated into basic schools (*Aufbauklassen*). This problem does not apply to Hamburg, Bremen, Rhineland-Palatinate, and Saarland. However, it might explain the striking ninth grade attendance between 1955-1960 in Hesse, where such classes were relatively common.

have left school after eight years otherwise, which applies directly to basic track students. However, the reform might have also affected the potential dropouts from the other tracks as afterwards, they had to comply to at least $C9$ grades of schooling, as well.

3 Empirical strategy

Following the prior literature, I exploit the regional variation in the timing of schooling law changes within the 2SLS framework, which estimates the following outcome equation

$$y_{ist} = \alpha \mathit{Educ}_{ist} + \chi_s + \delta_t + X'_{ist}\beta + \epsilon_{ist}, \tag{1}$$

where y is the labor market outcome of an individual i from state s in year t. $Educ$ represents an individual's years of schooling while χ and δ are vectors of state and year of birth fixed effects (FE), respectively. X comprises quadratic in age, an indicator for gender, and indicators for year of observation. The corresponding first-stage equation is

$$\mathit{Educ}_{ist} = \pi \mathit{Inst}_{st} + \lambda_s + \theta_t + X'_{ist}\phi + \nu_{ist} \tag{2}$$

where λ captures state FE and θ is a vector of year of birth FE. $Inst$ represents the instrument based on schooling laws that vary over time and across states. Following Pischke and von Wachter (2008), I pursue a specification where $Inst$ corresponds to a binary variable that indicates whether an individual was required to attend nine instead of eighth years of compulsory schooling ($C9$). However, I also account for the simultaneous introduction of the short school years in several states given their potentially opposing effect on schooling duration and wages. To this purpose, X in both equations includes an indicator for being affected by the short school years (SSY). In a separate specification, I also explore using SSY as an additional instrument. I expect a positive first-stage coefficient on $C9$ and a negative on SSY.

The internal validity of the 2SLS approach rests on the assumptions that the instrument is relevant and exogenous. I test the relevance condition in Section 5.1. Exogeneity is fulfilled if all other changes that occur across states prior to reform are uncorrelated with the law change itself and the outcomes. The coefficient on the instrument is basically identified within a difference-in-differences framework as the empirical strategy involves both state and year of birth FE. Thus any factors that disproportionately affected states over time would violate the common trends assumption and lead to biased estimates. To account for such differences in trends across states, I estimate alternative specifications that differ by the set of covariates included in X: First, I add time-variant state-specific controls measured when an individual was 14 years old such as student-teacher ratio and average weekly wages, which should capture the underlying trends in schooling quality and earnings. Second, I explore inclusion of year of birth indicators that differ across more broadly defined geographical regions as suggested by Stephens and Yang (2014). Finally, I include state-specific trends in year of birth as in the main specification by Pischke and von Wachter (2008).

Regarding inference, I follow the original study and the majority of prior literature using state-specific schooling law instruments, which typically assumes that standard errors are correlated among individuals from the same state and birth cohort. However, the appropriate level of clustering in designs relying on variation across regions and over time is not straightforward. The conventional approach is to be conservative and to cluster at the most aggregate level feasible subject to finite sample issues (for a recent survey see Cameron and Miller (2015)). However, along the discussion in the latest installment in that debate by Abadie et al. (2017), the clustering within state/year of birth cells is justified by the quasi-experimental design because the treatment assignment is perfectly correlated within these clusters. Nevertheless, my main conclusions hold also when I use alternative inference approaches such as clustering by state, year of birth, the two-way clustering by state and year of birth, or Wild cluster bootstrap (Cameron et al., 2008) at the state level. Generally, the 2SLS confidence intervals would be incorrect in case of weak instruments (Staiger and Stock, 1997), but for the majority of my analysis, the first-stage F-Statistic testing the significance of the instrument is above the conventional threshold.[8]

4 Data

The empirical analysis draws on two main data sets. Following Pischke and von Wachter (2008), my primary data source is the Qualification and Career Survey (QaC), which is a repeated cross section of the German labor force above age 15 (GESIS, 2017). So far, there are six survey waves collected in 1979, 1985/86, 1991/92, 1998/99, 2005/06, and 2011/12. Each of them samples on average about 25,000 workers.[9] I make several sample restrictions to be able to cleanly estimate the effect of schooling on wages. I restrict the data to German citizens living in the ten West German states (excluding Berlin). I focus on the prime ages from 25 through 55 to ensure that individuals completed their education and are not yet affected by

[8] Nevertheless, to be conservative, I also estimated 2SLS confidence intervals based on the conditional likelihood ratio (CLR) test by Moreira (2003), which has good power properties when instruments are weak and allows for clustering within each state/year of birth cell. The results were similar to those presented here.

[9] The original sample size varies across years between 20,000 and 35,000 thousand observations. The survey started with German citizens in the age group 15-65 living in West German states, but over time has been extended to foreigners, older workers, and East German states.

early-retirement programs. Unlike Pischke and von Wachter (2008), I consider only individuals born between 1945 and 1960. I do so for two reasons: first, they were born and enrolled in school after World War II and any war related shocks (e.g., interruptions in educational career, stress, food shortage early in life, absent father etc.) could confound the estimates. There is growing evidence on the long-lasting consequences of World War II on socioeconomic outcomes. For example, Kesternich et al. (2014) find that wartime experiences decrease the number of completed school grades, but prolong the time needed to reach a given level of education. Second, the earlier cohorts were affected by the temporary extensions of compulsory schooling in the early 1950s. The predominantly local character of these laws hampers their parameterization because there is no systematic information on their implementation at the local level and my data do not provide a municipality identifier. Simply ignoring these temporary laws introduces a measurement error as we would assign the affected individuals the eight-year standard while they actually experienced extended schooling. A measurement error in the instrument should not cause a bias in the 2SLS estimate as it proportionally attenuates the first-stage and reduced-form coefficients, but it might yield noisier estimates.

The main dependent variable is log hourly wage calculated from the gross monthly earnings (in DM) and the actual number of working hours per week.[10] Given that the schooling system is structured by tracks rather than by highest attended grade, German data do not typically report the number of years of schooling. Nevertheless, the data include information on the highest degree attained, which basically corresponds to the completed track. I exclude individuals with specific school degrees that could have been obtained only in the socialist East Germany

[10] See Pischke and von Wachter (2008) for details on the original earnings variable in the QaC and the calculation of the hourly wage.

(GDR). The main advantage of QaC is however that it additionally asked respondents for the year of graduation from secondary school. Subtracting from this variable the year of birth and the typical age of school enrollment (six) gives the approximate number of years attended to primary and secondary school.[11] I restrict the sample to individuals having from six to 22 years of schooling to reduce the measurement error in the calculated schooling duration.[12]

To construct the schooling law indicators $C9$ and SSY, ideally, I would need the exact date of birth to determine the year of school enrollment, the direct information on the state where an individual went to school, and the attended school track. Unfortunately, I cannot use month of birth as the question was not included throughout. I thus assume that all individuals started school in the calendar year when they turned six years old, as illustrated in Figure 1, though this creates some mismeasurement.[13] For example, I assume that the Bavarian birth cohort 1954 was enrolled entirely in 1960 and thus not affected by $C9$. However, a substantial fraction of this cohort regularly started school one year later in 1961 and those individuals were already treated by $C9$. To minimize the measurement error in the instrument and to increase

---

[11] In absence of this measure, German studies typically construct years of schooling by imputing the usual duration of a particular track. However, this approach would require using the instrument variable $C9$ to determine whether an individual with a basic track degree should have graduated after 8 or 9 years. While Pischke and von Wachter (2008) provide alternative estimates based on the imputed schooling years, I abstain from this approach as it mechanically generates a strong first-stage relationship between $C9$ and $Educ$.

[12] The originally calculated number of years of schooling was between -57 to 48. The implausibly low and high values potentially reflect misreporting in the year of graduation or birth. My sample restriction corresponds to dropping the observations below the first and above the 99 percentile of schooling duration, but alternative cuts yield similar results. Pischke and von Wachter (2008, 2005) do not mention how they handle this issue.

[13] The exact cut-off date for school enrollment varies by state. For the analyzed cohorts, most states stuck to 31st March, but children born between 1st April and 30th June could also start school in the calendar year of their sixth birthday if they were sufficiently mature.

precision of the 2SLS estimates, I exclude the "pivotal", i.e., not yet fully affected cohorts from the main estimation sample. By doing so, I also follow the precautionary recommendations of the recent literature pointing to observations located at the threshold as a potential source of bias in regression discontinuity (RD) designs.[14]

Which cohorts in the affected states were exposed to the SSY depended on the attended school track at that time of the policy change (Pischke, 2007). However, I observe only the completed degree and it might be endogenous to the policy changes (as shown in Section 5.1 for the $C9$ reform and in Pischke (2007) for the SSY). Thus, I do not condition the coding on the completed track and assign the SSY as if everyone attended the lowest track (as documented in Figure 1). Thus, for example, all students born between 1953 and 1961 experienced SSY because they attended school in April 1966. In Schleswig-Holstein, Bremen, and Saarland, the SSY affected also the cohort 1952 as, under the $C9$ regime, it entered their last grade in April 1966. The marginal cohorts 1952/1953 and 1961 experienced one SSY and those in between two SSY. I do not apply this distinction in my main analysis, but Section 6 shows that an alternative definition yields identical results.

The state of school attendance is reported only in 2006. For consistency and comparison with earlier results, I use the current state of residence as a proxy throughout. The resulting measurement error should be limited as regional mobility in Germany is generally low and

[14] Some authors point to heaping and manipulation concerns (e.g., Barreca et al., 2016), others to problems arising by construction (e.g., Fort et al., 2016). I do not apply a standard RD design, but I do pool (stack) multiple thresholds generated by the staggered implementation of the C9 reform across the states. Thus, similar challenges might potentially occur in my identification strategy. Because I do not observe the exact date (or month) of birth, I cannot explicitly investigate to what extend and for what reason the "pivotal" cohorts might be problematic.

about 77% of individuals in the 2006 sample currently live in the state of their school attendance. This share is 87% for basic track graduates, who were mostly affected by the $C9$ reform. Nevertheless, in Section 6, I demonstrate that using the available information on state of school attendance leads to similar results.

The second data source are individual register records from the Sample of Integrated Labour Market Biographies (Antoni et al., 2016).[15] I use the SIAB data primarily to investigate potential effects on employment because QaC samples only working individuals. The SIAB is a 2% sample of population covered at least once by the social security system between 1975 and 2014 due to employment or benefit receipt. Since 2000, registered jobseekers with no benefit eligibility and participants of employment or training measures are also included. The original data cover about 80% of the total workforce in Germany as civil servants and self-employed are not subject to social security (though some might be included e.g., as voluntarily registered jobseekers). The SIAB is organized by spells and follows the sampled individuals as long as their activities appear in social security records. To efficiently handle the data volume, I reshape the spells into a yearly panel with 1st July being the day of observation, but alternative cutoffs (e.g., 1st January, 1st April, 1st October) yield the same results.

Similar to my QaC sample, I focus on German citizens born 1945-1960 and their outcomes measured at ages 25-55. I exclude a small number (0.3%) of apprentices, trainees, and interns, who contribute to social security while being in education. Given that the state of residence is only available since 1999 and even then frequently not reported, I impute the missing values

15 Specifically, I use the weakly anonymous version 1975-2014 and accessed the data via a Scientific Use File at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) in Nuremberg.

with the location of the current unemployment agency or employer.[16] Unfortunately, the information on educational attainment is limited to school leaving certificate and it lumps together the basic and intermediate degree. Consequently, I am not able to construct any measure of years of schooling and focus on reduced-form effects. In contrast to QaC, the social security records do not facilitate identification and exclusion of individuals who went to school in the GDR. To reduce the measurement error in the instrument, I restrict my sample to individuals who entered the data before the fall of Berlin Wall (9th November 1989). Earnings are reported as gross daily wage (in EUR) with no information on working hours. The different earnings measure would limit comparisons with the QaC results only if the $C9$ reform affected working hours, which in not the case (see Section 5.4). I convert the original values into DM.

The SIAB data have several advantages over the Micro Census (MC) used as a complementary data set by Pischke and von Wachter (2008). First, it allows a larger sample because the scientific use files of the MC cover 0.7% of households (as opposed to 2% of workforce in SIAB). Second, SIAB is based on employers' payroll records, so that the information on gross earnings is likely very accurate. In contrast, there is no earnings measure in the MC; respondents report only their monthly net income, which comprises any income sources including labor, public transfers, properties, and support from family members. Relating all these to working hours yields a very inaccurate proxy for hourly wages. Finally, the SIAB is a panel encompassing nearly entire occupational careers of the cohorts I study while the MC is a rotating panel with a quarter of the sample being replaced each year. However, the MC data do not include a personal identifier, so that most individuals enter the pooled MC sample in Pischke

[16] Generally, about 94% of employees in Germany work and live in the same state (DESTATIS, 2017). The results remain very similar when I alternatively use the location of the first employer recorded in the SIAB panel.

and von Wachter (2008) repeatedly and the authors cannot adjust for multiple occurrences in their estimations.[17]

My final QaC sample comprises 35,250 individuals and the SIAB panel about 4.1 million annual observations on 209,137 individuals. Because for a given person the instrument is time-invariant, I cannot use the panel structure of the register data. Given the age restrictions, an individual might appear up to 31 times in the 40-year panel. The median incidence is 27 times. To avoid repeated occurrence in the estimations, I randomly select one observation per person, but I also run pooled regressions using all observations and cluster the standard errors by individual. Alternatively, while pooling all observations, I weight them by the inverse of the number of times each individual appears in the panel, which gives each person an equal weight in the estimations. I link both data sets to time-variant state-specific controls such as average weekly wages and student-teacher ratio from administrative data (STBA, 2006). The aggregate variables are measured as of the year when an individual was 14 years old, i.e., in the eighth school year. Table A.1 provides summary statistics.

5 Results

5.1 Effects of the schooling laws on educational attainment

To investigate the first-stage relationship between the endogenous years of schooling and the instrument, I estimate equation 2 using the QaC data. Table 2 shows the results from four different model specifications. I display only the coefficient on the instrumental variable $C9$ and the newly included variable SSY capturing the effect of short school years. In addition,

[17] Another advantage is that SIAB is free of charge. Acquiring the MC for the same period (1975-2014) generates a substantial cost because the current fee is 250 EUR per survey year.

all regressions include a constant, gender dummy, quadratic in age, and the maximum set of indicators for state, year of birth, and survey year.

I start with the standard specification without any further covariates. Column 1 yields that the extension of compulsory schooling from eight to nine years ($C9$) increased the duration of schooling by almost 0.4 years. The point estimate is statistically significant and the first-stage F-statistic exceeds the conventional weak instrument threshold of 10. The coefficient on SSY also shows the expected direction and implies that the short school years compressed the length of schooling by 0.25 years. Adjusting for state-specific developments in wages and school quality in column 2 slightly increases the magnitude of both coefficients and the F-Statistic. To mitigate the concern that the state-specific controls measured at age 14 are endogenous to the law changes, I also explored using a different timing of the data linkage and obtained nearly identical results.[18]

Alternatively, the regression in column 3 adds year of birth FE that differ across broadly specified regions because Figure 1 suggests an existence of a regional gradient in the timing of the $C9$ reform. I distinguish between north (Schleswig-Holstein, Hamburg, Lower Saxony, and Bremen), middle (North Rhine-Westphalia, Hesse, Rhineland Palatinate, and Saarland), and south (Baden-Wurrtemberg and Bavaria) Germany.[19] This specification flexibly accounts for any kind of time-varying factors that are common for the states within each region. The point coefficient on $C9$ decreases only slightly. The larger standard error and the lower F-Statistic are

[18] The earliest possible linkage is at age 12 because previous statistics on school quality are not available for some states. Merging at age 13 or 15 also generates similar results.

[19] Generally, aggregating the West German states into broader regions seems arbitrary, but alternative definitions yield similar results; I also explored a split into north versus south and into regions corresponding to the three occupational zones after World War II: American, British, and French.

not surprising given that the number of states within each region is small.

Finally, column 4 includes state-specific trends in year of birth,[20] which yields the strongest first stage. A worrying feature of this specification is that it might run the risk that the trends pick up not only different preexisting developments across states, but rather confound the long-run developments with dynamic responses to a law change (Wolfers, 2006). Reassuringly, the estimates for the effect of $C9$ on schooling duration are relatively stable across the different specifications and statistically indistinguishable from each other. The F-Statistics confirm the relevance of the instrument throughout.

Using the last specification in Table 2, Pischke and von Wachter (2008) repeat the estimations on a subsample limited to basic track students and find lower estimates than for the entire sample. They interpret this finding as supportive for the conclusion of zero returns because basic track students were mostly affected by the extension of compulsory schooling, so that we might expect that any effects of $C9$ are larger among this group compared to the full sample. However, the authors emphasize that this analysis rests on the assumption that the reform did not affect the track choice itself. Indeed, they find a small and insignificant effect on the probability of completing the basic track. In Table 3, I re-investigate this issue by using my sample and specification to estimate the reduced-form effects of $C9$ on completed degree. The dependent variable in each column corresponds to one out of five mutually exclusive degrees.

Overall, the estimates imply that $C9$ decreased the fraction of school dropouts and lead to substantial shifts between the basic and the intermediate track. The increased incidence of intermediate degree is not surprising given its lower opportunity cost after the reform; the

[20] Following Pischke and von Wachter (2008), I focus on specifications using linear trends, but Section 6 shows also estimates adjusted for quadratic trends.

difference between compulsory schooling duration and time spent in the middle track shrunk from two to one year. In additional estimations, I found that this shift toward the middle track is very robust to various changes in the model specification such as dropping the state-specific trends, omitting the SSY indicator, and including the year of birth FE that vary across regions. Given the endogenous selection into track, limiting the sample to basic track students loses its validity and would produce biased estimates.

5.2 Effect of years of schooling on hourly wages

Having established that the $C9$ reform significantly affected the duration of schooling, I now turn to the 2SLS estimates of the effect of increased schooling on wages. Panel A of Table 4 shows the results from the four different model specifications corresponding to the first-stage estimates from Table 2. The instrumental variable is $C9$, but I control for SSY throughout. Again, I assume that standard errors are correlated among individuals from the same state and birth cohort, but inference results based on other assumptions (see Table A.2) generally lead to similar conclusions. Panel B documents the corresponding reduced-form estimates of the direct effect of both law changes on wages. All effects are identified via a differences-in-differences design, because state and year of birth FE are also included. Panel C displays the conventional OLS estimates.

Given that the dependent variable is log hourly wage, the estimate in column 1 of Panel A corresponds to a 9% wage return per additional year of schooling. However, this result relies on the strong assumption that there are no other factors that correlate with the treatment and disproportionately affect states over time. To make this assumption more credible, in column 2, I control for state-specific developments in wages and school quality, which leads to a slightly stronger effect. The model in column 3 replaces the state-specific controls by interaction terms between regions and year of birth FE. The 2SLS effect is comparable to the previous ones, but less precisely estimated. Finally, the specification in column 4 includes state-specific linear trends in year of birth, which yields the most conservative effect. Nevertheless, the coefficient corresponds to a wage return of 8% and remains statistically significant at the 1%-level.

In international comparison, wage returns of this magnitude are in line with the majority of literature that documents rather moderate wage effects from extensions in compulsory schooling in Europe (see e.g., Meghir and Palme, 2005; Brunello et al., 2009; Aakvik et al., 2010; Dolton and Sandi, 2017). Still, my sizable and significant estimates challenge the previous study by Pischke and von Wachter (2008), who conclude about zero returns to schooling in Germany by using the $C9$ reform, the same data, and a similar model specification to the last column in Table 4. However, I introduced some important modifications and the next section discusses their role in explaining the contradictory findings.

5.3 Comparison with Pischke and von Wachter (2008)

I start the comparison by showing how undoing the major changes compared to Pischke and von Wachter (2008) affects my baseline findings. Table 5 summarizes the results. For comparison, column 1 repeats my main estimate. I focus on the impact of three essential modifications: omitting birth cohorts prior to 1945, excluding the "pivotal" cohorts, and controlling for the parallel introduction of the SSY in some states.[21]

[21] I also revised the timing of the $C9$ reform for four states and additionally included the two newest survey waves of the QaC. However, the two changes turn to be of a negligible importance as I obtain a virtually identical 2SLS estimate of 0.081 when I use the coding of $C9$ found in Pischke and von Wachter (2008) and the survey years available back then. Detailed results are available upon request.

First, I excluded cohorts 1930-1944 for two reasons: the temporary extensions of compulsory schooling in the early 1950s (see Figure 1) and war-related shocks (e.g., disruptions in schooling). Regarding the first, these cohorts might have already experienced prolonged education, so that assigning them to the eight year regime would create a mismeasurement in the instrument. While a measurement error in $C9$ should not cause a bias in the 2SLS estimate, it might yield noisier estimates. Regarding the wartime turbulences, they might lead to bias if such shocks are not captured by cohort FE. It seems quite likely that the war affected various states differently, and my restricted sample mitigates the risk of such confounding effects. After including the problematic cohorts in column 2, both the first-stage and the reduced-form coefficients in Panels B and C are considerably weaker though the reduced-form estimate decreases relatively more. The patterns point rather to the confounding wartime effects than to a potential impact of the measurement error. The 2SLS estimate in Panel A loses its precision and decreases in magnitude, but is statistically indistinguishable from the baseline result. In Section 6, I demonstrate that my main findings are robust to various changes in the range of included birth cohorts as the estimated returns remain always in range of 6%-8%. Nevertheless, my preferred cohort selection 1945-1960 yields the most efficient estimate.

Second, I also excluded the "pivotal" cohorts only partly affected by the $C9$ reform due to the cutoff rules for school enrollment to reduce a measurement error and to avoid any potential bias caused by the observations at the threshold (e.g., Barreca et al., 2016; Fort et al., 2016). Again, including them in column 3 does not vanish my main result. Indeed, the first-stage and reduced-form estimates are lower in magnitude and less precise, but the 2SLS coefficient still allows to conclude about significant and economically important wage returns to schooling. However, the uncertainty of the inference is now much larger comparable to my baseline results.

Finally, I modified the empirical strategy by including SSY as an additional control. Thus column 4 displays the results after omitting SSY. The 2SLS estimate decreases in magnitude compared to column 1, but remains significant. I argue that including SSY appears inevitable given its parallel introduction with the $C9$ reform in several states and the offsetting effects of both legislative changes in the first-stage and reduced-form regressions (see Tables 2 and 4).

Generally, the three minor but plausible adjustments help to cleanly identify the effect of interest, but neither of them fully accounts for the contradictory findings by itself. Thus rather their combination appears to be crucial for establishing sizable and precisely estimated returns to schooling in this framework. Column 5 shows the results obtained after removing all three modifications. The first-stage and reduced-form coefficients almost exactly match the respective estimates of 0.190 and 0.010 reported by Pischke and von Wachter (2008) in their Table 1 (column 1) and Table 2 (column 3).[22] The small and imprecise 2SLS estimate does no longer allow to conclude about significant wage returns to schooling.

5.4 Effects of the ninth compulsory school year on employment and related outcomes

An important question is whether the $C9$ reform affected labor force participation because a different sorting into employment threatens a causal interpretation of the estimated wage returns. It is likely that employers prefer more educated job applicants over those who spent less time in

[22] The article prints a reduced-form estimate of -0.010, but the negative sign is presumably a typing error because earlier Pischke and von Wachter (2005) report 0.010. The negligible differences between their estimates and my replication mostly reflect the revised timing of the $C9$ reform and inclusion of the two most recent waves of QaC, which I keep in Table 5 for comparability with my main results. A remaining divergence occurs because Pischke and von Wachter (2005, 2008) do not mention how they handle some important data issues such as, e.g., the implausibly low and high numbers of schooling years, which I mention in Section 4.

school because of expected productivity improvements.[23] To investigate potential employment effects, I now turn to the register data including both employed and unemployed individuals.

Table 6 shows the reduced-form estimates for the effect of the $C9$ reform on various outcomes.[24] I cannot exploit the panel structure of the data to estimate the coefficient of interest within a FE framework because there is no within-person variation in the instrument. Thus, the table displays results from three different approaches. In Panel A, I randomly select one observation per person to avoid repeated occurrence in the estimations. In Panel B, I run the regressions on pooled data and cluster the standard errors by individual. In Panel C, I also pool all observations, but weight them by the inverse of the number of times each individual appears in the panel data.

The dependent variable in column 1 is an indicator for being employed, which is equal to one if a person has any positive labor income subject to social security contributions. The coefficient on $C9$ is in expected direction, but small and insignificant. Similarly, I do not find any effects on the probability of having a parallel secondary job in column 2. The precise zero effects on

[23] Unfortunately, I am not able to provide any strong evidence that the $C9$ reform actually led to human capital improvements due to lack of comprehensive data on this issue for the cohorts under study. While following Kamhöfer and Schmitz (2016), I investigated the potential cognitive gains using the SOEP, I was left with less than 2,000 observations and the estimates were too imprecise and sensitive to various model specifications to draw any conclusions.

[24] The corresponding first-stage coefficient obtained from the QaC data is 0.507 (see Table 2, column 4), so that inflating the reduced-form coefficients by two (1/0.507≈2) yields approximately the two-sample 2SLS estimates. Such two-sample 2SLS estimates might be interpreted only approximately as the QaC generates five and the SIAB 39 dummies for the year of observation. This implies that the vector of covariates in the first-stage and reduced-form regressions is slightly different. Moreover, both stages are disproportionately affected by various sources of measurement error, so that its impact does not cancel in the cancel in the two-sample 2SLS estimates.

labor supply are in line with no significant changes in eligibility for public transfers such as unemployment benefits and welfare in column 3, though employment and benefit receipt are not necessarily mutually exclusive. The results in all three panels lead to identical conclusions.

In the last column, I repeat the analysis for log daily wages. All panels yield significant impacts of the reform on this earnings measure. However, the magnitudes of these effects do not exactly match the reduced-form estimate of 0.042 from the QaC data (see Table 4, Panel B, column 4) and they also differ across the Panels A through C.

The generally weaker reduced-form effects in the SIAB compared to the QaC data are not driven by the different definition of the outcome (daily earnings vs. hourly wages) as the QaC yields a nearly identical reduced-form effect on both measures (see Table A.3). This is not surprising as I do not find any corresponding effects of the $C9$ reform on working hours. Two other reasons are more likely to explain the discrepancy in the reduced-form effects from the different data sets. First, to some extent, the divergence reflects a different sample composition as SIAB records do not cover self-employed and civil servants. Indeed, excluding these groups from my QaC sample decreases the reduced-form estimate to 0.032, which is close to the estimate in Panel A. Section 6 documents that this change is driven by self-employed and not by civil servants. Second, the weaker reduced-form estimates in the SIAB might be also partly due to the limited regional information. In this respect, the QaC is superior because it surveys the state of residence, which is missing for about 65% of my SIAB sample. I then use the state of work, but it might lead to a larger mismeasurement in the instrument, thereby attenuating the reduced form disproportionately stronger compared to the QaC.[25]

[25] I tested this argument in the QaC data by replacing the state of residence by state of work, which is available in waves 2006 and 2012. Indeed, the reduced-form estimate decreases from 0.042 to 0.038. This change

Nevertheless, these two reasons do not explain the difference between the reduced-form effects on earnings across the three panels in Table 6. To understand this issue, I investigate the compositional differences between the different SIAB samples. I focus here on the differences between Panels A and B as the point estimates for Panel C lie in between. I find a higher average age and a lower percentage of women if I pool all observations as in Panel B (see Table A.1). Figure A.1 plots the age structure in the data. The dotted black line shows that the SIAB in Panel B lacks young observations below age 30. This shortage is per construction because the records start in 1975, so that I do not observe individuals born between 1945 and 1950 at particular ages between 25 and 29. Similarly, I don't observe the birth cohort 1960 at age 55 as the data end in 2014. In contrast, the dashed line for the SIAB sample in Panel A uncovers an over-representation of individuals below age 33. This pattern is mainly driven by women and might reflect that some of them leave the labor force due to childbearing, thereby being more likely to enter the sample at a young age when I randomly draw one observation per person. To make the age structure in the SIAB samples more comparable and assure that it is not driven by the period of data availability, I restrict the data to earnings measured between the 30th and 54th age year and find that the estimates in all panels converge (see Table A.4).

The finding that the magnitude of the effect of $C9$ might depend on the age composition in the data suggests that monetary returns to compulsory schooling are not necessarily constant over the entire occupational career. While a detailed investigation of the potentially different effects from a life course perspective extends beyond the scope of this paper, to shed some light on the question when the effect materializes, I re-run the regressions for (log) daily earnings by

is generated by a relatively small subsample because the two waves constitute only 13% of my entire QaC sample.

age groups (see Table A.5). The effects appear only for relatively experienced workers.

Overall, the complementary analysis using social security records confirms a significant impact of the ninth compulsory school year on earnings. The magnitude of the reduced-form effects is lower compared to the estimate from the QaC data mainly because of a different sample composition and a larger measurement error in the $C9$ variable. I do not find any evidence that the monetary returns are driven by an endogenous selection into employment. Interestingly, it seems to matter when the returns are measured as they might vary over occupational careers.

6 Robustness

This section assesses the sensitivity of my main findings from Table 4 to alternative model specifications and sample restrictions. Table 7 summarizes the results by showing the 2SLS coefficient for the effect of one extra school year on wages and the corresponding reduced-form and first-stage results. All regressions include state-specific trends in year of birth and an indicator for individuals affected by the SSY. For comparison, I repeat the baseline results at the top of the table.

A major threat to the identification strategy is that the estimates are generated by differences between states as opposed to variation within states over time (Stephens and Yang, 2014). To rule out the possibility that any remaining regional differences explain the results, I step-wise excluded every single state from the analysis and found that the results remained remarkably stable. Panels A through C exemplify this robustness when omitting selected states, which could actually confound the estimates for specific institutional reasons: in the city states Hamburg and Bremen, the tracking took place two grades later compared to the remaining states. Hesse introduced $C9$ gradually across counties as described in Section 2. In Lower Saxony, the SSY

did not affect the basic track students, but shortened the schooling duration in the two higher tracks (Pischke, 2007). Generally, Panels A through C confirm my main conclusions.

In panel D, I depart from the main specification by adding quadratic state-specific trends in year of birth. The wage return is still significant, but its magnitude decreases to 6%. This change is mainly driven by the remarkably stronger first stage because the reduced-form estimate remains nearly unaffected. The purpose of including squared trends is to account for slow-moving trends in each state prior to the reform. While this appears to be a rigorous strategy to satisfy the common trends assumption, the approach is also controversial because flexible trends can inadvertently pick up reform-induced dynamics and produce biased results (see e.g., Wolfers, 2006; Lundborg et al., 2014). Consequently, the 2SLS estimate in the first column may not be the closest to the true value.

I next test whether the results are driven by a potential measurement error in the instrument and the schooling measure. The main source of mismeasurement in $C9$ is the limited geographical information. The QaC 2006 survey is an exception because in addition to the current state of residence, it also reports the actual state of school attendance. Panel E shows that using this information does not appreciably change the estimates.[26] Panel F intends to further reduce the potential measurement error in the endogenous schooling variable by omitting individuals who obtained less than eight and more than 13 years of schooling. Before the $C9$ reform, these numbers corresponded to the expected schooling duration in the lowest and the highest track, respectively. This sample cut is very conservative as some deviations from the typical dura-

[26] Unfortunately, is not possible to use only the QaC 2006 to directly compare estimates based on state of residence versus state of schooling because only individuals born between 1951 and 1960 are observed at ages 25-55 in 2006. This considerably reduces not only the sample size but also the variation in the instrument (see Figure 1), which leads to a weak fist stage.

tion are possible e.g., due to repeating grades, skipping a year, or switching between tracks. Nevertheless, Panel F yields even a higher 2SLS estimate for the wage return.

The next two panels focus on alternative definitions of the SSY measure. First, in Panel G, I replace the dummy variable SSY by two indicators to account for separate effects of experiencing one versus two SSY. The results are identical. To exploit the entire variation in the SSY, the variable should be ideally coded depending on the attended school track prior to the policy change, which is unfortunately not reported. Using the available information on the highest completed degree is not innocuous because it was endogenously affected by the $C9$ reform (as shown in Section 5.1). Panel H illustrates the consequences. At first glance, the estimated wage returns in the first column are still in line with the baseline results, though slightly smaller in magnitude. However, the reduced-form and first-stage estimates turn their direction, which is implausible. Although not shown, the respective effects of the SSY also follow dubious patterns, which was also found for a similar specification by Pischke (2007).[27] These alternative estimates support my argument that conditioning on completed track is highly problematic given that the track is already an outcome of the treatment.

Panels I and J exclude civil servants and self-employed, respectively, who might potentially face different wage rigidity compared to conventionally employed workers. While omitting civil servants leaves the results unchanged, excluding self-employed reduces the wage returns to

[27] In his main specification, Pischke (2007) additionally includes dummies for completed track although he shows that exposure to SSY affected track choice. By doing so, he finds a small negative and insignificant reduced-form effect of the SSY on wages in the QaC data. I can replicate this result for SSY when I control for completed track. However, in addition to the endogeneity issue, such specification seems awkward when looking at wage returns to schooling because completed track largely determines years of schooling. Nevertheless, by doing so, I obtain a significant 2SLS estimate of 0.078, which is close to my baseline result, though less precise as the standard error is 0.046.

5.5%. This finding suggests that this group benefits relatively more from an extra schooling year than remaining workers or that extended schooling affects the incidence of self-employment. Unfortunately, auxiliary regressions yielded only fragile evidence for a different sorting into self-employment due to the schooling extensions because the coefficient on $C9$ was imprecisely estimated and sensitive to specification changes. Thus, I cannot entirely exclude that the $C9$ reform lead to some compositional changes of the work force, so that omitting self-employed in Panel I and Table 6 might induce a selected sample.

Generally, any effects of the reform on labor labor market participation might bias the wage returns. However, given that I did not find any employment effects in Section 5.4, selective labor supply is not a concern. To further strengthen the argument, I also re-run the main regressions separately for women and men in Panels K and L, respectively. The rationale for doing so is that selective labor force participation should be less an issue among men while extended schooling might affect women's labor supply through potential fertility effects (Cygan-Rehm and Maeder, 2013).[28] The first-stage results imply that the reform more strongly affected women's education compared to men. The reduced-form coefficients indicate also slightly higher wage responses among women, though they lack precision due to reduced sample sizes. Overall, the magnitude of the 2SLS estimates for the wage returns in the first column is very similar across gender.

I also carefully tested to what extent my selection of included birth cohorts affects the results. For example, I widened the sample to earlier cohorts by using different birth years as cut-offs. Figure A.2 in the appendix illustrates that the estimated returns slightly vary depend-

[28] Indeed, using the female subsample, I found significantly negative effect of schooling on the number of children living in a woman's household of -0.2. The effect is slightly higher compared to Cygan-Rehm and Maeder (2013), who look at completed fertility and do not account for the SSY. The corresponding coefficient for men is positive (0.084) and insignificant.

ing on the cohort selection, but remain always in range of 6%-8% and statistically significant at the 5% level, at least. The figure uncovers also that my preferred cohort selection 1945-1960 yields the most precise result, but the estimate obtained for cohorts 1936-1960 is nearly identical. Generally, all displayed estimates are statistically indistinguishable from each other as the confidence intervals overlap throughout. In Panel M of Table 7, I further limit the period to cohorts 1946-1959, which also leads to similar conclusions.

Finally, Panel N shows results from an alternative specification that includes SSY as an additional instrument instead of treating it as an ordinary covariate. Table 2 documents that both law changes significantly affected schooling duration. Moreover, the offsetting effects of $C9$ and SSY are of similar magnitude. However, the interventions differently dealt with the curriculum; while the $C9$ introduced an additional year of instruction, the SSY compressed the old curriculum into a shorter instructional period. Therefore, using them both as instruments implies that the identification of the local average treatment effect (LATE) comes from a different group of compliers compared to my baseline specification. Nevertheless, Panel N strongly supports my main conclusions as the 2SLS estimate virtually matches my main result.

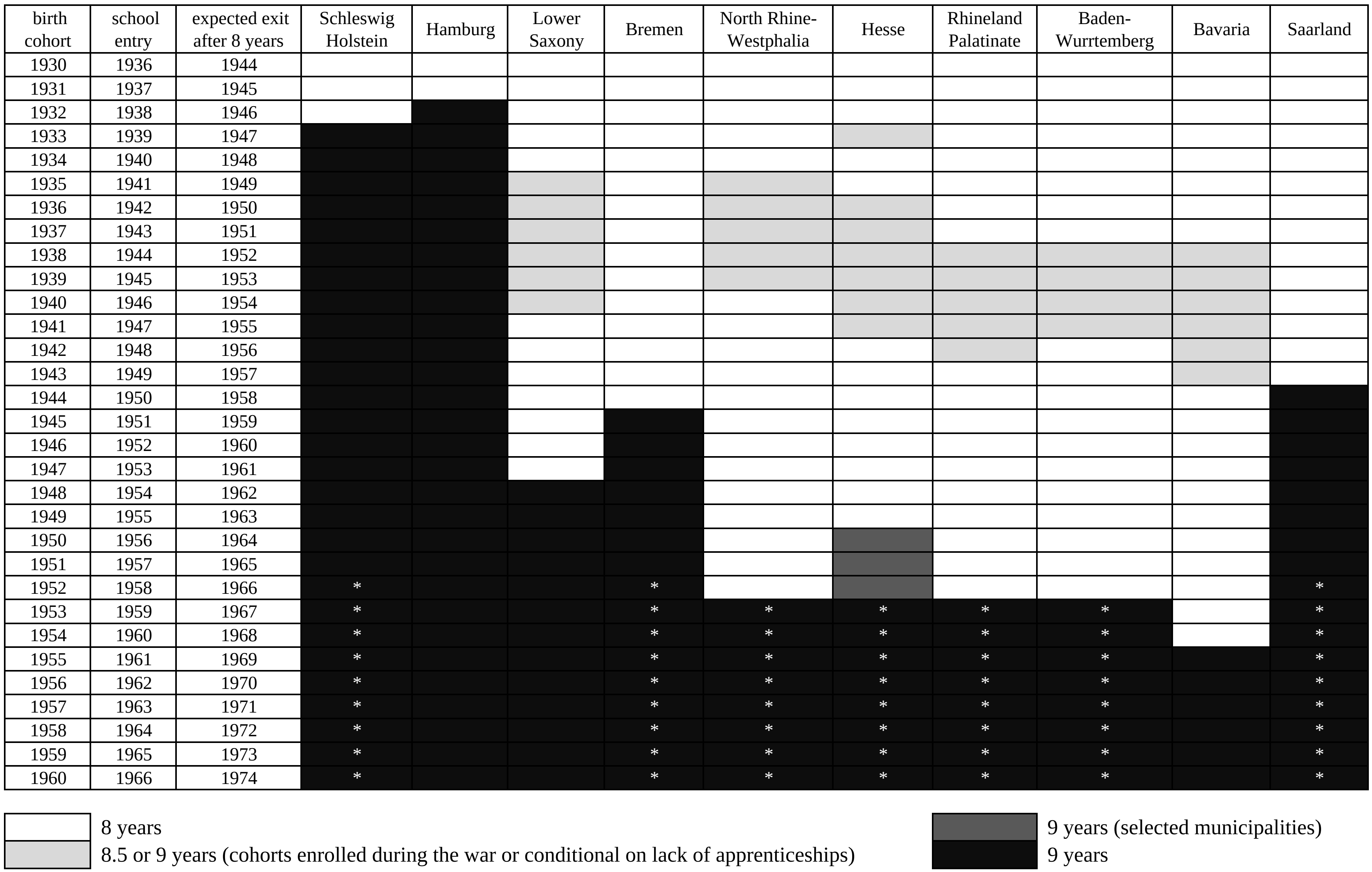
7 Conclusions

The existence of important economic returns to compulsory schooling is one of the most investigated and controversial topics within the economics literature. The empirical evidence from studies that seriously address the endogeneity of schooling to establish causality is largely inconclusive. For example, for several years, the general consensus among researchers using U.S. data was that an additional year of schooling increases individual earnings by 6%-10% (see e.g., Card, 1999). Only recently, Stephens and Yang (2014) presented an appealing alternative to the earlier conclusions; they obtain insignificant and even wrong-signed estimates after introducing a minor modification to a common empirical strategy using U.S. schooling laws as instruments. One of the first studies that obtained a zero return from extensions of compulsory schooling was Pischke and von Wachter (2008) for Germany. The authors argued that German students acquire the labor market-relevant skills much earlier compared to students in the U.S. or other countries. So far, this conclusion has been widely accepted and cited.

This paper provides new evidence for the existence of sizable monetary returns to compulsory schooling in Germany. Specifically, I find wage effects of 6-8% per additional year of schooling by implementing minor though essential modifications to the sample and specification from the earlier study. My adjustments are mainly for institutional reasons and I show that each of them alone does not eliminate my estimate of the return to schooling. It is rather a combination of unconsidered institutional details that alter the previous estimates by Pischke and von Wachter (2008). My baseline results are not confounded by other policy changes and not driven by preexisting trends across states as highlighted in Stephens and Yang (2014). Moreover, the estimated returns remain robust in various sensitivity tests and are also detectable in alternative data from social security records. Looking beyond earnings, I do not find any benefits of extended schooling in terms of employability and independence of public transfers.

The finding that compulsory schooling positively affects wages in Germany yields new implications for the latest iteration of the general debate, especially if such returns are nonexistent in the U.S. The established argument that general secondary education in Germany carries less value added in terms of skills compared to the U.S. has lost its validity. Unfortunately, I am not able to shed more light on potential mechanisms behind the estimated returns due to lack of comprehensive data on cognitive skills for the cohorts under study and leave the issue for future research. Nevertheless, my estimates may help to explain the puzzle of why the German extensions of compulsory schooling lead to significant responses along other dimensions such as e.g., health (Kemptner et al., 2011) and fertility (Cygan-Rehm and Maeder, 2013) if the wage returns were apparently zero. This paper vindicates the importance of the income channel as a potential mechanism behind such non-pecuniary benefits of schooling and also yields implications for future research exploiting this reform for identification. Generally, my findings call also for attention to institutional specifics within the quasi-experimental research as undervalued details might lead to very different conclusions and policy implications.

Figure 1: Duration of compulsory schooling over time and by state



*Cohorts affected by short school years (SSY) in basic track (April 1966 - November 1966, December 1966-July 1967)

Note: Own illustration based on Leschinsky and Roeder (1980), Petzold (1981), original legislation from the state laws, and school statistics from STBA (2006). Further details available on request.

Table 1: Ninth graders in the basic track relative to eight graders previous year

	Federal state									
Year	Schl.-Holstein	Hamburg	Lower Saxony	Bremen	NRW	Hesse	Rhinel.-Palat.	Baden-Wurtt.	Bavaria	Saar-land
1953	31%	N/A	15%	N/A	15%	4%	93%	3%	0%	N/A
1954	46%	N/A	5%	N/A	2%	4%	91%	3%	1%	N/A
1955	65%	N/A	2%	N/A	0%	8%	89%	3%	1%	N/A
1956	72%	N/A	3%	N/A	1%	8%	90%	2%	1%	N/A
1957	72%	N/A	3%	N/A	1%	10%	24%	3%	1%	0%
1958	75%	N/A	5%	N/A	1%	13%	1%	2%	1%	99%
1959	75%	N/A	4%	N/A	1%	14%	2%	7%	1%	87%
1960	74%	78%	4%	73%	1%	13%	1%	5%	1%	86%
1961	79%	84%	2%	80%	1%	2%	0%	2%	0%	84%
1962	76%	82%	82%	77%	1%	4%	0%	2%	0%	76%
1963	78%	82%	75%	79%	2%	16%	1%	4%	0%	88%
1964	77%	82%	79%	80%	2%	37%	1%	4%	0%	82%
1965	77%	82%	78%	80%	4%	49%	1%	7%	0%	80%
Year of final introduction of the ninth compulsory school year										
from Leschinsky and Roeder (1980)										
	1947	1946	1962	1959	1967	1967	1967	1967	1969	1958
from Pischke and von Wachter (2008)										
	1956	1949	1962	1958	1967	1967	1967	1967	1969	1964

Source: NRW refers to North Rhine-Westphalia. Own calculations based on absolute numbers of students in their eight and ninth school year from STBA (2006).

Table 2: First stage: effect of the schooling laws on years of schooling

	(1)	(2)	(3)	(4)
$C9$	0.363***	0.383***	0.341***	0.507***
	(0.055)	(0.053)	(0.077)	(0.060)
F-Statistic	43.95	52.20	19.52	70.19
SSY	-0.244***	-0.285***	-0.327***	-0.433***
	(0.060)	(0.058)	(0.057)	(0.101)
State-specific controls	no	yes	no	no
Region $\times$ year of birth FE	no	no	yes	no
State $\times$ linear trend in year of birth	no	no	no	yes
Observations	33,250			
Clusters	154			

Notes: Sample restricted to (West-)German citizens born 1945-1960. The dependent variable is years of (primary and secondary) schooling calculated as year of graduation minus year of birth minus six. $C9$ and SSY are indicators for being affected by a ninth compulsory school year and short school years, respectively. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic). State-specific controls comprise weekly wages of male workers and student-teacher ratio measured at the age 14. Region distinguishes between north (Schleswig-Holstein, Hamburg, Lower Saxony, Bremen), middle (North Rhine-Westphalia, Hesse, Rhineland Palatinate, Saarland), and south (Baden-Wurrtemberg, Bavaria) Germany. Robust standard errors clustered at state $\times$ year of birth cells in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations. State-specific controls from STBA (2006).

Table 3: Reduced-form effects of the ninth compulsory school year on completed school degree

	(1) No degree	(2) Basic track	(3) Middle track	(4) High school	(5) Other
$C9$	-0.006 **	-0.049 ***	0.042 ***	0.015	-0.001
	(0.003)	(0.017)	(0.016)	(0.011)	(0.002)
Observations			33,250		
Clusters			154		

Notes: Sample restricted to (West-)German citizens born 1945-1960. The five outcome variables (in columns) correspond to indicators for a completed school degree. $C9$ is an indicator for being affected by a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, interview year, gender, and control for age (quadratic), state-specific linear trends in year of birth, and an indicator for being affected by short school years (SSY). Robust standard errors clustered at state $\times$ year of birth cells in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations.

Table 4: Effect of years of schooling on log wages

	(1)	(2)	(3)	(4)
Panel A: 2SLS				
Years of schooling	0.092 ***	0.099 ***	0.086 **	0.083 ***
	(0.026)	(0.029)	(0.035)	(0.026)
First-stage F-Statistic	43.95	52.20	19.52	70.19
Panel B: Reduced form				
$C9$	0.033 ***	0.038 ***	0.029 **	0.042 ***
	(0.008)	(0.011)	(0.012)	(0.013)
SSY	-0.044 ***	-0.046 ***	-0.046 ***	-0.050 ***
	(0.008)	(0.010)	(0.008)	(0.018)
Panel C: OLS				
Years of schooling	0.035 ***	0.035 ***	0.035 ***	0.035 ***
	(0.001)	(0.001)	(0.001)	(0.001)
State-specific controls	no	yes	no	no
Region $\times$ year of birth FE	no	no	yes	no
State $\times$ linear trend in year of birth	no	no	no	yes
Observations	33,250			
Clusters	154			

Notes: Sample restricted to (West-)German citizens born 1945-1960. The dependent variable is log hourly wage. $C9$ and SSY are indicators for being affected by a ninth compulsory school year and short school years, respectively. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic). State-specific controls comprise weekly wages of male workers and student-teacher ratio measured at the age 14. Region distinguishes between north (Schleswig-Holstein, Hamburg, Lower Saxony, Bremen), middle (North Rhine-Westphalia, Hesse, Rhineland Palatinate, Saarland), and south (Baden- Wurrtemberg, Bavaria) Germany. Robust standard errors clustered at state $\times$ year of birth cells in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations. State-specific controls from STBA (2006).

Table 5: Differences in sample and specification compared to Pischke and von Wachter (2008)

	(1)	(2)	(3)	(4)	(5)
Panel A: 2SLS					
Years of schooling	0.083 ***	0.065 **	0.060 *	0.068 **	0.039
	(0.026)	(0.027)	(0.031)	(0.027)	(0.037)
Panel B: First stage					
$C9$	0.507 ***	0.313 ***	0.390 ***	0.353 ***	0.195 ***
	(0.060)	(0.059)	(0.078)	(0.057)	(0.047)
F-Statistic	70.19	27.87	25.23	38.53	17.06
Panel C: Reduced form					
$C9$	0.042 ***	0.020 **	0.023 *	0.024 **	0.008
	(0.013)	(0.009)	(0.013)	(0.010)	(0.008)
Cohorts 1930-1944 excluded	yes	no	yes	yes	no
"Pivotal" cohorts excluded	yes	yes	no	yes	no
SSY included	yes	yes	yes	no	no
Observations	33,250	52,853	35,508	33,250	55,224
Clusters	154	300	160	154	310

Notes: Sample restricted to (West-)German citizens born 1930-1960. The dependent variable is log hourly wage. $C9$ and SSY are indicators for being affected by a ninth compulsory school year and short school years, respectively. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic) and state-specific linear trends in year of birth. Robust standard errors clustered at state $\times$ year of birth cells in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations.

Table 6: Reduced-form effects of the ninth compulsory school year on various labor market outcomes

	(1) Working	(2) Secondary job	(3) Public transfers	(4) Log daily earnings
Panel A: One randomly selected observation per person				
$C9$	0.003	0.001	-0.005	0.029 ***
	(0.005)	(0.002)	(0.004)	(0.008)
Observations	209,137	209,137	209,137	190,242
Clusters		154		
Panel B: All observations pooled				
$C9$	0.003	0.002	-0.003	0.015 **
	(0.002)	(0.001)	(0.002)	(0.007)
Observations	4,110,248	4,110,248	4,110,248	3,820,906
Clusters	209,137	209,137	209,137	190,242
Panel C: All observations pooled and weighted				
$C9$	0.002	0.002	-0.004	0.020 **
	(0.003)	(0.001)	(0.003)	(0.008)
Observations	4,110,248	4,110,248	4,110,248	3,820,906
Clusters	209,137	209,137	209,137	190,242

Notes: Sample restricted to (West-)German citizens born 1945-1960. $C9$ is an indicator for being affected by a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the introduction of short school years (SSY). Robust standard errors in parentheses. Standard errors are clustered at state $\times$ year of birth cells in Panel A and individual level in Panels B and C. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: SIAB 1975-2014, own calculations.

Table 7: Sensitivity analysis: 2SLS estimates of the effect of schooling on log wages

	(1) Coeff. on years of schooling	(2) Reduced form $C9$	(3) First stage $C9$	(4) First stage F-Stat.	(5) Observations [Clusters]
Baseline	0.083 ***	0.042 ***	0.507 ***	70.19	32,250 [154]
	(0.026)	(0.013)	(0.060)		
A: Exclude Hamburg & Bremen	0.079 ***	0.038 ***	0.474 ***	60.90	31,908 [122]
	(0.028)	(0.013)	(0.061)		
B: Exclude Hesse	0.087 ***	0.044 ***	0.511 ***	71.42	30,249 [139]
	(0.024)	(0.012)	(0.060)		
C: Exclude Lower-Saxony	0.074 **	0.039 ***	0.533 ***	40.05	29,271 [139]
	(0.028)	(0.014)	(0.084)		
D: Add quadratic trends	0.060 ***	0.041 ***	0.678 ***	91.60	33,250 [154]
	(0.021)	(0.014)	(0.071)		
E: Use state of schooling	0.078 ***	0.041 ***	0.521 ***	69.95	33,252 [154]
	(0.025)	(0.013)	(0.062)		
F: Schooling from 8 to 13 years	0.092 ***	0.040 ***	0.435 ***	45.88	28,751 [154]
	(0.035)	(0.013)	(0.064)		
G: Two indicators for SSY	0.083 ***	0.044 ***	0.526 ***	65.62	33,250 [154]
	(0.028)	(0.015)	(0.065)		
H: Coding of SSY based on	0.061 ***	-0.044 ***	-0.724 ***	12.93	33,250 [154]
completed track (degree)	(0.014)	(0.016)	(0.201)		
I: Exclude civil servants	0.082 ***	0.047 ***	0.571 ***	58.47	29,643 [154]
	(0.024)	(0.013)	(0.075)		
J: Exclude self-employed	0.055 **	0.028 **	0.510 ***	59.00	30,735 [154]
	(0.022)	(0.012)	(0.066)		
K: Women	0.080 *	0.047	0.590 ***	30.53	13,352 [154]
	(0.048)	(0.030)	(0.107)		
L: Men	0.088 *	0.041 **	0.458 ***	27.87	19,898 [154]
	(0.045)	(0.019)	(0.087)		
M: Cohorts 1946-1959	0.074 ***	0.039 ***	0.526 ***	61.67	28,999 [134]
	(0.026)	(0.014)	(0.067)		
N: Two instruments: $C9$ & SSY	0.084 ***	0.042 ***	0.507 ***	35.11	33,250 [154]
	(0.027)	(0.013)	(0.060)		

Notes: Sample restricted to (West-)German citizens born 1945-1960. The dependent variable is log hourly wage. $C9$ and SSY are indicators for being affected by a ninth compulsory school year and short school years, respectively. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the SSY. Robust standard errors clustered at state $\times$ year of birth cells in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations.

Appendix

Table A.1: Sample means

	QaC	SIAB one observation per person		SIAB all observations pooled	
Working	1.00	0.92	1.00	0.93	1.00
Hourly wage (log)	2.88	N/A	N/A	N/A	N/A
Daily wage (log)	N/A	4.59	4.59	4.73	4.73
Year of observation	1,991.19	1,993.12	1,992.94	1,994.12	1,993.95
Year of birth	1,953.02	1,953.34	1,953.28	1,953.51	1,953.47
Age (years)	38.17	39.78	39.65	40.61	40.48
Female	0.40	0.47	0.47	0.43	0.42
Affected by $C9$	0.64	0.66	0.65	0.67	0.67
Affected by SSY	0.39	0.41	0.40	0.42	0.41
Years of schooling	10.58	N/A	N/A	N/A	N/A
Highest degree: no	0.01				
Highest degree: basic	0.51	} 0.70	0.75	0.75	0.79
Highest degree: intermediate	0.28				
Highest degree: high school	0.20	0.13	0.14	0.11	0.12
Highest degree: other/n.a.	0.00	0.17	0.12	0.14	0.09
Schleswig-Holstein	0.05	0.04	0.04	0.04	0.04
Hamburg	0.03	0.03	0.03	0.03	0.03
Lower Saxony	0.12	0.12	0.12	0.12	0.12
Bremen	0.01	0.01	0.01	0.01	0.01
North Rhine-Westphalia	0.29	0.28	0.27	0.28	0.27
Hesse	0.09	0.10	0.10	0.09	0.09
Rhineland Palatinate	0.07	0.06	0.06	0.06	0.06
Baden-Wurrtemberg	0.14	0.16	0.16	0.16	0.16
Bavaria	0.18	0.18	0.19	0.19	0.19
Saarland	0.02	0.02	0.02	0.02	0.02
State-specific weekly wages (DM)	247.04	253.14	252.05	256.14	255.26
State-specific student-teacher ratio	34.08	33.86	33.85	33.79	33.79
Secondary job	N/A	0.02	0.03	0.02	0.03
Public transfers eligibility	N/A	0.08	0.01	0.06	0.01
Observations	33,250	209,137	190,242	4,110,248	3,820,906

Notes: Sample restricted to (West-)German citizens born 1945-1960. $C9$ and SSY are indicators for being affected by a ninth compulsory school year and short school years, respectively. State-specific variables as of age 14 from STBA (2006).
Source: QaC 1979-2012 and SIAB 1975-2014.

Table A.2: Alternative inference results

	(1)	(2)	(3)	(4)
2SLS point estimate on years of schooling	0.092	0.099	0.086	0.083
Panel A: Cluster robust standard errors				
Baseline: clustering by state × year of birth (g=154)	(0.026)	(0.029)	(0.035)	(0.026)
Two-way clustering by state & year of birth (g_1=10, g_2=16)	(0.013)	(0.016)	(0.016)	(0.013)
Clustering by year of birth (g=16)	(0.021)	(0.024)	(0.024)	(0.022)
Clustering by state (g=10)	(0.019)	(0.022)	(0.031)	(0.020)
Wild restricted efficient bootstrap clustering by state (g=10)	(0.026)	(0.036)	(0.079)	(0.031)
Panel B: 95% confidence intervals				
Baseline: clustering by state × year of birth (g=154)	[0.042; 0.142]	[0.043; 0.156]	[0.017; 0.155]	[0.032; 0.134]
Two-way clustering by state & year of birth (g_1=10, g_2=16)	[0.067; 0.117]	[0.068; 0.131]	[0.054; 0.117]	[0.057; 0.109]
Clustering by year of birth (g=16)	[0.051; 0.133]	[0.052; 0.146]	[0.039; 0.132]	[0.040; 0.126]
Clustering by state (g=10)	[0.054; 0.130]	[0.055; 0.143]	[0.026; 0.146]	[0.045; 0.121]
Wild restricted efficient bootstrap clustering by state (g=10)	[0.025; 0.129]	[0.028; 0.168]	[0.008; 0.319]	[-0.005; 0.118]
Panel C: p-values for the null hypothesis				
Baseline: clustering by state × year of birth (g=154)	0.000	0.001	0.015	0.001
Two-way clustering by state & year of birth (g_1=10, g_2=16)	0.000	0.000	0.000	0.000
Clustering by year of birth (g=16)	0.000	0.000	0.000	0.000
Clustering by state (g=10)	0.000	0.000	0.005	0.000
Wild restricted efficient bootstrap clustering by state (g=10)	0.027	0.018	0.028	0.063

Notes: g refers to the number of clusters. The Wild restricted efficient bootstrap (Davidson and MacKinnon, 2010) extends the Wild cluster bootstrap (Cameron et al., 2008) to instrumental variables estimators. I apply the recommended bootstrap procedure with 999 replications, imposing the null hypothesis, and using the Rademacher weights. Its immediate result is the p-value for the null hypothesis in Panel C. The resulting confidence intervals in Panel B are nonsymmetric. Panel A shows conservative estimates of the standard errors calculated as the width of a corresponding 95% confidence interval divided by 2 × 1.96 (Cameron and Miller, 2015).
Source: QaC 1979-2012, own calculations.

Table A.3: Reduced-form effects of the ninth compulsory school year on daily wages and working hours

	(1) Log hourly wages (baseline)	(2) Log daily wages	(3) Working hours	(4) Log working hours
$C9$	0.042 ***	0.041 ***	-0.290	-0.001
	(0.013)	(0.013)	(0.329)	(0.009)
Observations		33,250		
Clusters		154		

Notes: Sample restricted to (West-)German citizens born 1945-1960. $C9$ is an indicator for being affected by a ninth compulsory school year. Daily wages correspond to gross monthly earnings divided by 20 working days. Working hours are reported per week. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the introduction of short school years (SSY). Robust standard errors in parentheses. Standard errors are clustered at state $\times$ year of birth cells. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: QaC 1979-2012, own calculations.

Table A.4: Reduced-form effects of the ninth compulsory school year on various labor market outcomes measured at ages 30 - 54

	(1) Working	(2) Secondary job	(3) Public transfers	(4) Log daily earnings
Panel A: One randomly selected observation per person				
$C9$	0.004	0.000	-0.007	0.024 ***
	(0.005)	(0.002)	(0.005)	(0.009)
Observations	198,987	198,987	198,987	180,398
Clusters		154		
Panel B: All observations pooled				
$C9$	0.002	0.001	-0.003	0.017 **
	(0.002)	(0.001)	(0.002)	(0.008)
Observations	3,438,508	3,438,508	3,438,508	3,195,753
Clusters	198,987	198,987	198,987	180,398
Panel C: All observations pooled and weighted				
$C9$	0.002	0.001	-0.005	0.020 **
	(0.003)	(0.001)	(0.003)	(0.009)
Observations	3,438,508	3,438,508	3,438,508	3,195,753
Clusters	198,987	198,987	198,987	180,398

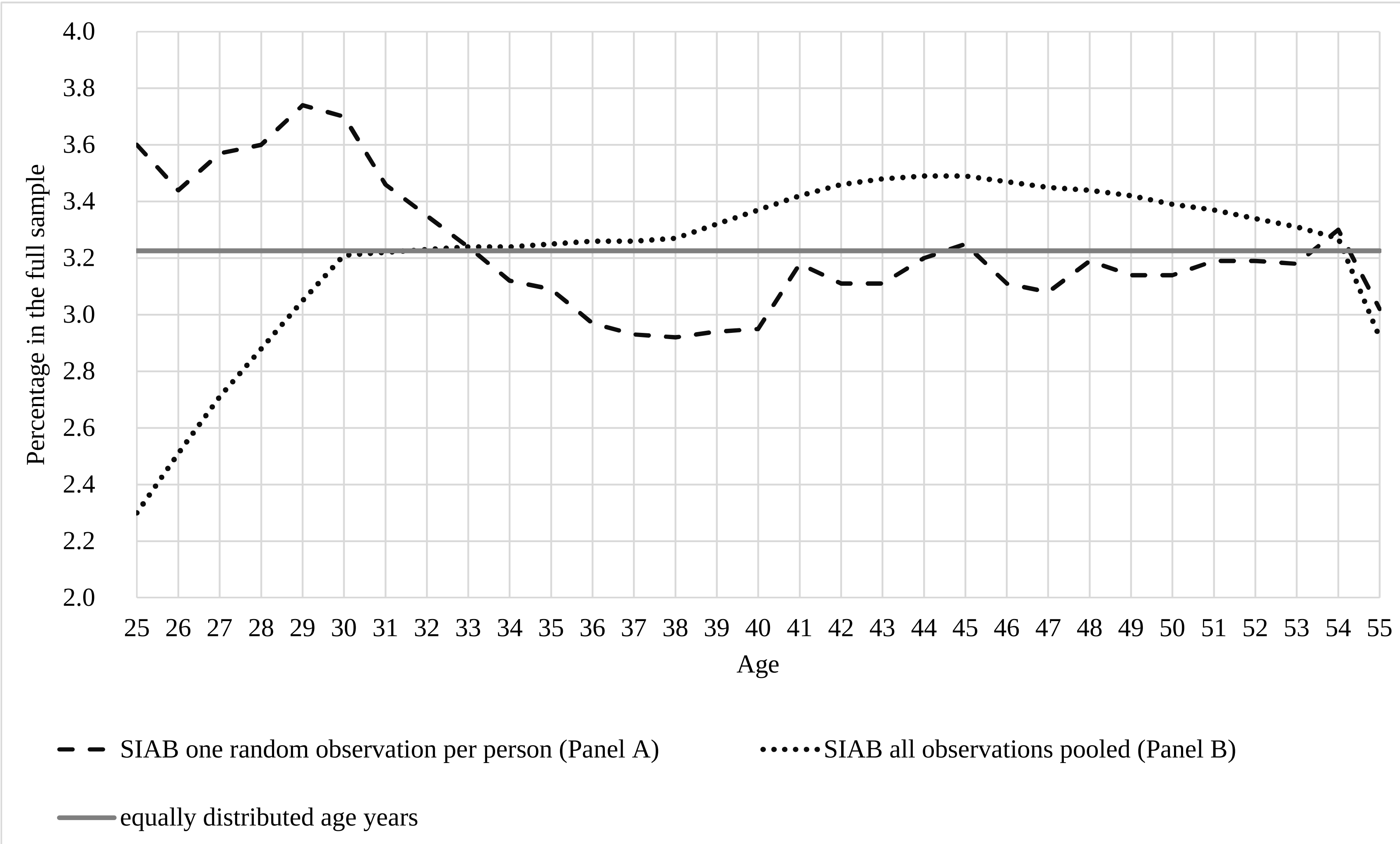
Notes: Sample restricted to (West-)German citizens born 1945-1960. $C9$ is an indicator for being affected by a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the introduction of short school years (SSY). Robust standard errors in parentheses. Standard errors are clustered at state $\times$ year of birth cells in Panel A and individual level in Panels B and C. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: SIAB 1975-2014, own calculations.

Table A.5: Reduced-form effects of the ninth compulsory school year on daily earnings by age

	(1)	(2)	(3)	(4)	(5)
Age group	30-34	35-39	40-44	45-49	50-54
Panel A: One randomly selected observation per person					
$C9$	0.008	0.007	0.007	0.029 **	0.033 **
	(0.006)	(0.007)	(0.007)	(0.013)	(0.014)
Observations	147,336	143,350	147,191	144,044	138,734
Clusters	154	154	154	154	154
Panel B: All observations pooled					
$C9$	0.007	0.006	0.014 *	0.023 **	0.032 **
	(0.007)	(0.008)	(0.009)	(0.011)	(0.013)
Observations	616,796	624,654	657,750	659,733	636,820
Clusters	147,336	143,350	147,191	144,044	138,734
Panel C: All observations pooled and weighted					
$C9$	0.005	0.005	0.008	0.022 *	0.040 ***
	(0.008)	(0.008)	(0.009)	(0.012)	(0.015)
Observations	616,796	624,654	657,750	659,733	636,820
Clusters	147,336	143,350	147,191	144,044	138,734

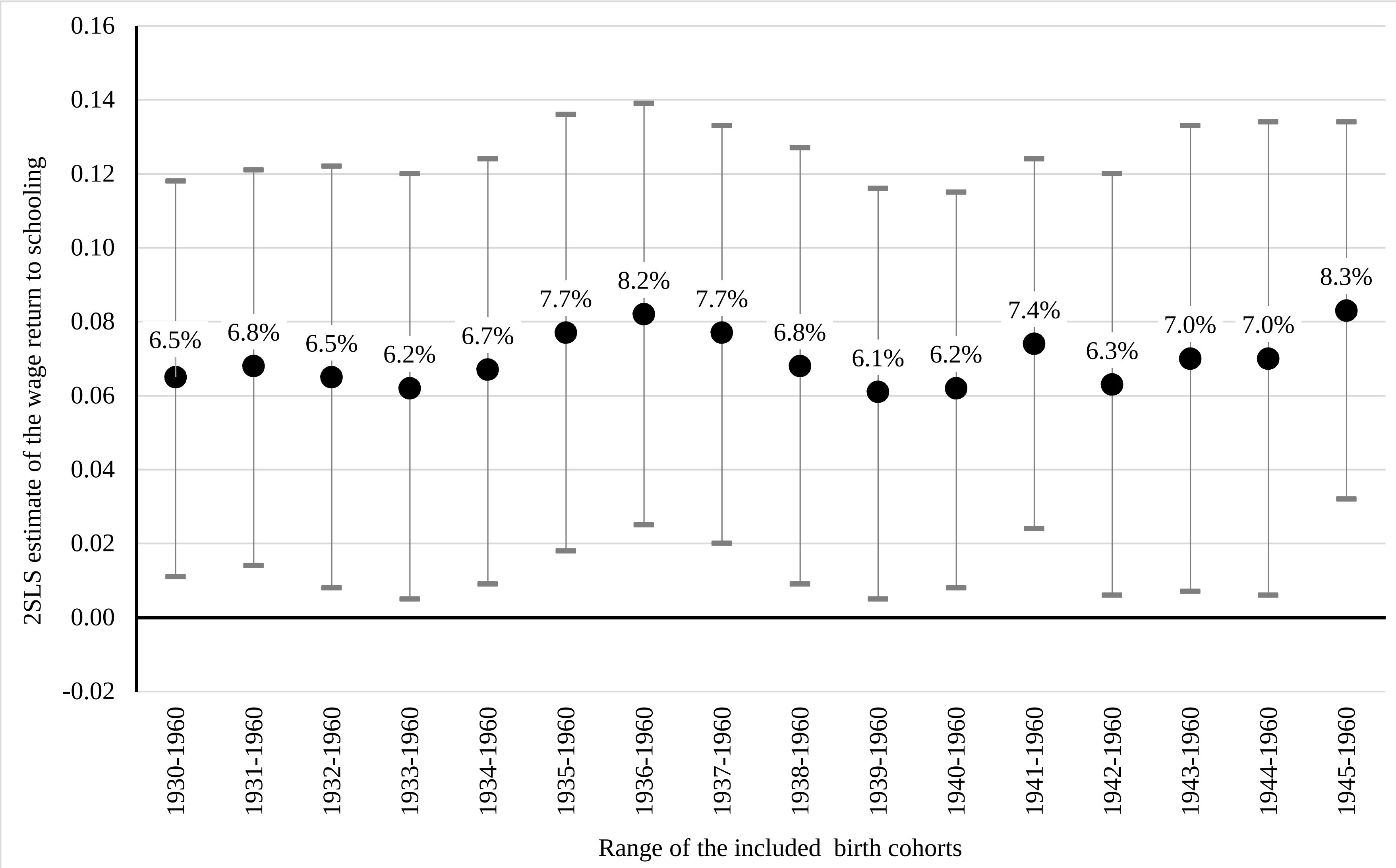
Notes: Sample restricted to (West-)German citizens born 1945-1960. The dependent variable is log daily earnings. $C9$ is an indicator for being affected by a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the introduction of short school years (SSY). Robust standard errors in parentheses. Standard errors are clustered at state $\times$ year of birth cells in Panel A and individual level in Panels B and C. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level.
Source: SIAB 1975-2014, own calculations.

Figure A.1: Age structure in the SIAB data



Source: SIAB 1975-2014.

Figure A.2: Effect of years of schooling on log wages across the range of included birth cohorts



Notes: Each dot shows a 2SLS point estimate of the effect of years of schooling on log wages and the corresponding 95% confidence interval. All regressions include a constant, indicators for state, year of birth, year of interview/observation, gender, and control for age (quadratic), state-specific linear trends in year of birth, and the introduction of short school years (SSY). Robust standard errors are clustered at state $\times$ year of birth cells. All samples are restricted to (West-)German citizens.
Source: QaC 1979-2012, own calculations.

References

Aakvik, Arild, Kjell G. Salvanes, and Kjell Vaage (2010). Measuring heterogeneity in the returns to education using an education reform. *European Economic Review 54*(4), 483–500.

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017). When should you adjust standard errors for clustering? NBER Working Paper 24003, National Bureau of Economic Research (NBER), Cambridge.

Acemoglu, Daron and Joshua D. Angrist (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual 15*, 9–59.

Angrist, Joshua D. and Alan B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics 106*(4), 979–1014.

Antoni, Manfred, Andreas Ganzer, and Philipp vom Berge (2016). Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2014. FDZ data report, 04/2016 (en), Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nürnberg.

Barreca, Alan I, Jason M Lindo, and Glen R Waddell (2016). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry 54*(1), 268–293.

Becker, Gary S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy 70*(5), 9–49.

Brunello, Giorgio, Margherita Fort, and Guglielmo Weber (2009). Changes in compulsory schooling, education and the distribution of wages in Europe. *Economic Journal 119*(536), 516–539.

Cameron, A Colin and Douglas L Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Card, David (1995). Using geographic variation in college proximity to estimate the returns to schooling, on aspects of labour market behaviour. In L. N. Christofides, K. Grant, and R. Swidinsky (Eds.), *Aspects of labour market behaviour: Essays in honour of John Vanderkamp*, pp. 201–222. Toronto: University of Toronto Press.

Card, David (1999). Chapter 30 - The Causal Effect of Education on Earnings. Volume 3, Part A of *Handbook of Labor Economics*, pp. 1801 – 1863. Amsterdam: Elsevier.

Cygan-Rehm, Kamila and Miriam Maeder (2013). The effect of education on fertility: evidence from a compulsory schooling reform. *Labour Economics 25*, 35–48.

Davidson, Russell and James G MacKinnon (2010). Wild bootstrap tests for iv regression. *Journal of Business & Economic Statistics 28*(1), 128–144.

DESTATIS (2017). Bevölkerung und Erwerbstätigkeit 2016. Erwerbsbeteiligung der Bevölkerung. Ergebnisse des Mikrozensus zum Arbeitsmarkt. Fachserie 1 Reihe 4.1, Statistisches Bundesamt (DESTATIS), Wiesbaden.

Devereux, Paul J. and Robert A. Hart (2010). Forced to be rich? Returns to compulsory schooling in Britain. *Economic Journal 120*(549), 1345–1364.

Dolton, Peter and Matteo Sandi (2017). Returning to returns: revisiting the British education evidence. *Labour Economics 48*, 87–104.

Fort, Margherita, Andrea Ichino, and Giulio Zanella (2016). On the perils of stacking thresholds in RD designs. Unpublished manuscript, available online at http://www.andreaichino.it/research_progress/fort_ichino_zanella_stacking.pdf [last assessed: 18.04.2018], European University Institute.

GESIS (2017). Qualifikation und Berufsverlauf 1979-2012 (Qualification and Career Survey 1979-2012). Survey description. Available online at https://dbk.gesis.org/dbksearch/ [Last assessed: 18.04.2018], Leibniz-Institut für Sozialwissenschaften (GESIS), Köln.

Grenet, Julien (2013). Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws. *Scandinavian Journal of Economics 115*(1), 176–210.

Harmon, Colm and Ian Walker (1995). Estimates of the economic return to schooling for the United Kingdom. *American Economic Review 85*(5), 1278–1286.

Kamhöfer, Daniel A. and Hendrik Schmitz (2016). Reanalyzing zero returns to education in Germany. *Journal of Applied Econometrics 31*(5), 865–872.

Kemptner, Daniel, Hendrik Jürges, and Steffen 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.

Kesternich, Iris, Bettina Siflinger, James P Smith, and Joachim K Winter (2014). The effects of World War II on economic and health outcomes across Europe. *Review of Economics and Statistics 96*(1), 103–118.

Leschinsky, Achim and Peter M. Roeder (1980). Didaktik und Unterricht in der Sekundarschule I seit 1950 - Entwicklung der Rahmenbedingungen. In Jürgen Baumert, Achim Leschinsky, Jens Naumann, Jürgen Raschert, and Peter Siewert (Eds.), *Bildung in der Bundesrepublik Deutschland - Daten und Analysen, Band 1: Entwicklungen seit 1950*, Chapter 4, pp. 283–392. Stuttgart: Klett-Cotta.

Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth (2014). Parental education and offspring outcomes: evidence from the swedish compulsory school reform. *American Economic Journal: Applied Economics 6*(1), 253–278.

Meghir, Costas and Mårten Palme (2005). Educational reform, ability, and family background. *American Economic Review 95*(1), 414–424.

Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.

Moreira, Marcelo J. (2003). A conditional likelihood ratio test for structural models. *Econometrica 71*(4), 1027–1048.

Oreopoulos, Philip (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review 96*(1), 152–175.

etzold, Hans-Joachim (1981). *Schulzeitverlängerung: Parkplatz oder Bildungschance? Die Funktion des 9. und 10. Schuljahres*. Bensheim: Päd.-Extra-Buchverlag.

Piopiunik, Marc (2014). Intergenerational transmission of education and mediating channels: Evidence from a compulsory schooling reform in germany. *Scandinavian Journal of Economics 116*(3), 878–907.

Pischke, Jörn-Steffen (2007). The impact of length of the school year on student performance and earnings: evidence from the German short school years. *Economic Journal 117*(523), 1216–1242.

Pischke, Jörn-Steffen and Till von Wachter (2005). Zero returns to compulsory schooling in Germany: evidence and interpretation. NBER Working Paper 11414, National Bureau of Economic Research (NBER), Cambridge.

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.

Schultz, Theodore W. (1961). Investment in human capital. *American Economic Review 51*(1), 1–17.

Staiger, Douglas and James H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

STBA (2006). Statistisches Jahrbuch für die Bundesrepublik Deutschland (Statistical Yearbook for the Federal Republic of Germany). Digital version - various years. Available online at http://resolver.sub.uni-goettingen.de/purl?PPN514402342 [Last assessed: 18.04.2018], Hrsg. Statistisches Bundesamt (STBA), Stuttgart: Metzler-Poeschel.

Stephens, Melvin and Dou-Yan Yang (2014). Compulsory education and the benefits of schooling. *American Economic Review 104*(6), 1777–1792.

Wolfers, Justin (2006). Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review 96*(5), 1802–1820.