IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14470

Are There No Wage Returns to Compulsory Schooling in Germany? A Reassessment

Kamila Cygan-Rehm

JUNE 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14470

Are There No Wage Returns to Compulsory Schooling in Germany? A Reassessment

Kamila Cygan-Rehm
University of Erlangen-Nürnberg, CESifo, IZA and LASER

JUNE 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Are There No Wage Returns to Compulsory Schooling in Germany? A Reassessment*

This study replicates and challenges the finding of zero wage returns to compulsory schooling in Germany by Pischke and von Wachter (Review of Economics and Statistics, 90(3) 2008, 592-598), which is unusual in the literature yet widely cited and until now uncontradicted. I document that this finding is sensitive to minor changes in sample restrictions and model specification. Further results suggest that their estimates are potentially confounded by previously unconsidered institutional details. These findings render the conclusion that compulsory schooling in Germany yields no wage returns at a minimum controversial.

JEL Classification: I21, I26, J31

Keywords: returns to schooling, education, wages, Germany, replication, reassessment

Corresponding author:
Kamila Cygan-Rehm
Friedrich-Alexander University Erlangen-Nürnberg
Dept. of Economics
Lange Gasse 20
90403 Nuremberg
Germany
E-mail: kamila.cygan-rehm@fau.de

* This paper uses proprietary data that can be obtained by filing a request with the Research Data Centers (FDZ) at the Federal Institute for Vocational Education and Training (BIBB) in Bonn. The author is willing to assist. I greatly appreciate helpful comments from the Editor Heather Anderson and three anonymous referees. I would like to thank Steve Pischke for sending me his STATA programs and providing extensive feedback on my early results. I also thank Colin Cameron, Deborah Cobb-Clark, Christina Gathmann, Mathias Hübener, Chris Karbownik, Daniel Kühnle, Regina T. Riphahn, Hendrik Schmitz, and participants in the annual LASER meeting 2017, ESPE 2018, the IZA World Labor Conference 2018, the VfS annual conference 2018, the CESifo Area Conference on Employment and Social Protection, SOLE 2019, and EALE 2019 for comments and advice. Lena Höcker provided excellent research assistance. I acknowledge financial support from Joachim Herz Stiftung (project no. 600047). The author declares no conflict of interest.

1 Introduction

The existence and the magnitude of wage returns to compulsory schooling is a much-disputed topic in economics. From a theoretical perspective, additional schooling is a human capital investment (Mincer, 1974), which is expected to pay off in terms of higher earnings. This prediction has been confirmed by extensive empirical research using quasi-experimental designs that aim at establishing causality (e.g., Card, 1999; Meghir and Palme, 2005; Brunello et al., 2009; Oreopoulos, 2006; Devereux and Hart, 2010; Bhuller et al., 2017). There are only a few exceptions in the literature presenting zero effects (Pischke and von Wachter, 2008; Grenet, 2013; Stephens and Yang, 2014).

One of the first studies that obtained zero returns to compulsory schooling was Pischke and von Wachter (2008) (henceforth PvW) for Germany. They exploit a staggered extension of compulsory schooling duration from eight to nine years ($C9$) across West German states after World War II. The unusual finding of zero wage returns in Germany has received considerable attention and has been widely cited by subsequent research. It has also been confirmed by Kamhöfer and Schmitz (2016) on different data.[1] Indeed, PvW document mostly statistically insignificant instrumental variable (IV) estimates; many of them are also economically small and some even negative. Is widely accepted to conservatively interpret such results as zero. However, in several specifications, they also obtain economically not-trivial effects, which is puzzling and argues for a reassessment.[2]

[1]They use a small cross-sectional sample from the German Socio-economic Panel (SOEP), and also find no effect on cognitive skills. Due to data limitations, their sample restrictions and model specification slightly deviate from PvW's setting; e.g., they exclude city-states Bremen and Hamburg and do not control for state-specific trends. A detailed discussion of the differences is beyond the scope of this paper. Nevertheless, the later study suffers from similar problems as those I argue to beset PvW's study.

[2]The lack of wage returns from the German schooling extensions is also surprising in light of evidence for significant non-pecuniary effects of this reform on health (Kemptner et al., 2011), fertility (Cygan-Rehm and Maeder,

This paper returns to the original data and results by PvW. Using the programs kindly provided by Steve Pischke, I can nearly exactly replicate their estimates. Going beyond a pure replication, I then probe their findings in a wide range of robustness tests. This analysis highlights that PvW's main estimate is very sensitive to minor changes in sample restrictions and model specification. I also show that several data choices made by PvW are both debatable and of consequence for their finding of zero returns. This casts doubts on whether the conclusion of zero wage returns to compulsory schooling in Germany holds. Further results suggest that PvW's estimates might be confounded by other schooling reforms such as the parallel implementation of short school years (SSY) in several states (see Pischke, 2007) and temporary schooling extensions preceding the actual passage of the final $C9$ laws.

More generally, this paper illustrates that public policies do not necessarily occur in isolation, which complicates their evaluation. My reassessment involved extensive background research and collecting additional data to uncover important institutional details (e.g., on the temporary extensions). This adds to the growing literature emphasizing that re-examinations of earlier research are important but may require more than merely replicating existing results (e.g., Ho et al., 2019).

2 Results

I start with a replication of PvW's main results using the Qualification and Career Survey (QaC).[3] A brief description of the data and empirical strategy is available in Online Appendix

2013), intergenerational transmission of education (Piopiunik, 2014), and even natives' attitudes towards immigrants, which persist beyond the treated generation (Margaryan et al., 2019). Related research for Germany also indicates positive wage returns to education beyond compulsory schooling (Kamhöfer et al., 2019).

[3]PvW also present an auxiliary evidence from the Micro Census (MC), which allows only for reduced-form analysis due to the lack of an appropriate educational measure. Thus, I focus on the main results from the QaC.

A. I show the original estimates in the top row of Table 1 and my replication directly below.[4] In all columns, the regression coefficients and standard errors are nearly identical. The remaining differences are in the third decimal place, and there is a small difference in sample size.[5]

Of major importance is the IV coefficient in column 1, which translates to wage returns of 5.8% per additional year of compulsory schooling, but is insignificant at conventional levels. Columns 2 and 3 show relevant details from the corresponding first-stage and reduced-form regressions.[6] PvW re-run the regressions on a subsample of basic track graduates, who were mostly affected by the reform. These estimates support their conclusion of zero returns. I show my replication results for this subsample in Tables C.1 and C.2 in Online Appendix C.

In Panel B of Table 1, I test whether PvW's main estimates are robust to plausible changes in sample restrictions and model specification. First, I assess to what extent the results are driven by a specific selection of the included birth cohorts. PvW limit their sample to individuals born between 1930 and 1960. Row S1 shows that starting with cohort 1931 increases the magnitude of the IV estimate and yields a statistically significant return of 6.7%. That the conclusions change after excluding one particular cohort is an extreme example but might simply be by chance. Thus, in S2, I further limit the sample to cohorts 1945-1960. Again, I obtain a significant and even larger return. While not displayed in the table, I also find significant returns in

[4]I was also able to accurately replicate PvW's results from other model specifications and their summary statistics reported in (Pischke and von Wachter, 2005), and this seems to be as close as a complete replication (see Table C.3 in Online Appendix C). The only striking difference is the percentage of self-employed in the sample, which is presumably due to a typo because PvW's code reveals that self-employed are included throughout.

[5]This presumably reflects minor revisions in the QaC data, which I received at least 12 years after PvW. Their version is no longer available, but the BIBB generously sent me the very first data files. The original data, however, do not fit the STATA code provided by Steve Pischke with respect to variable names and the coding of missing values. The initial sample sizes in the original files and the current version are identical. However, by comparing some key variables, I found minor disparities in the number of observations with missing values. Thus, it seems that over time, the BIBB reclassified some potentially implausible responses into missing values, which might explain the slightly smaller sample compared to PvW.

[6]PvW report a reduced-form estimate of -0.010, but the negative sign is a typo because the first-stage and the IV effects are both positive. Earlier, Pischke and von Wachter (2005) show a positive reduced-form effect of 0.010.

other cohort restrictions (e.g., 1934-1960, 1935-1960, 1936-1960 etc.).

For various reasons, data limited to more recent cohorts might produce different results. PvW's sample covers a long and unstable period of Germany's history and includes individuals born and attending school during World War II. Wartime experiences might have long-lasting consequences for educational careers (e.g. Kesternich et al., 2014). Any war-related shocks (e.g., interruptions in schooling) could confound PvW's estimates if the war affected various states differently, which seems likely. Unfortunately, it is difficult to directly test this hypothesis.

Another issue is that some cohorts in PvW's sample were exposed to temporary extensions of compulsory schooling in the early 1950s, which intended to address a high youth unemployment in the post-war period (see Appendix B for details). These statewide laws were gradually suspended as the labor market situation improved; in most affected states, at least seven years before the final $C9$ reform. PvW's instrument assigns individuals exposed to the temporary extensions to the control group, which is inappropriate because they potentially experienced more than eight years of compulsory schooling. Admittedly, these laws are difficult to parameterize because in many cases, municipalities could decide on the law's enforcement at a local level and the QaC data do not provide municipality identifiers. Nevertheless, by using the variation across states and over time, I code a crude indicator for exposure to the temporary extensions. I include it as an additional instrument in S3.[7] Alternatively, in S4, I use one combined instrument, which indicates the exposure to any extension (temporary or final). Technically, this is a sum of the two instruments used in S3. Both specifications yield non-zero wage returns.

Obviously, to some extent, the diverging results across various cohort selections might also

[7]The first-stage effect of the additional instrument on schooling is 0.073 and significant (standard error of 0.034). Given that the model is over-identified, I can also directly test the exogeneity assumption. The Sargan-Hansen tests for over-identifying restrictions yields a p-value of 0.203, which supports the validity of the instruments.

reflect a changed age composition in the data. Given that PvW use data collected between 1979 and 1999, by construction, they start to observe the early cohorts by the end of their forties and the more recent cohorts at the beginning of their occupational careers (see Table A.1 in Online Appendix A). Thus, excluding the early cohorts shifts the focus toward prime-age and younger workers. This might affect the estimates if wage returns to schooling vary over the life course (Bhuller et al., 2017) or if labor force participation at older ages is selective.

PvW study individuals between 19 and 65 years old. In S5, I drop the oldest workers aged above 60. In the German context, it might be desirable to even further restrict the sample due to early retirement programs. Thus, the sample in S6 comprises only individuals aged up to 55. In S7, I exclude the youngest workers, below age 25, which is the typical approach in the literature to mitigate the risk of endogenous sample selection.[8] PvW's conclusions do not hold in these other plausible age restrictions. This might suggest important heterogeneities. Figure C.1 in Online Appendix C shows further checks across 77 different age groups. The vast majority of the IV estimates are larger than PvW's and significant at the 10% level (86% of the cases and almost 50% of them at the 5% level).

In S8, I test to what extent the results are driven by outliers by excluding the 0.5% of the bottom and top wages each year. The extremely low and high values likely reflect a measurement error due to misreporting. Again, the IV estimate rises in magnitude and becomes significant.

Another relevant question is whether PvW's model was misspecified, meaning that the $C9$ indicator could be capturing other changes across states. A natural candidate for a potential confounder is the parallel implementation of the $SSYs$, which also affected schooling dura-

[8] For example, Oreopoulos (2006) includes the age group 25-64 in his U.S. and Canadian samples and 32-64 in the British data. Similar restrictions for various European countries apply Devereux and Hart (2010), Brunello et al. (2009), and Grenet (2013). Recently, for the U.S., Stephens and Yang (2014) focus on ages 25-54.

tion but in the opposite direction. Thus, in S9, I control for potential exposure to the $SSYs$.[9] In S10, I add squared state-specific trends to more flexibly capture all remaining factors that disproportionately affected states over time. Again, PvW's conclusions do not hold up.[10]

PvW argue that selective labor force participation could bias the results but it should be less of an issue among men than among women. In a footnote, they mention that they find significant returns when they restrict their MC sample to men but attribute this finding to sample variability. In S11, I re-examine this issue with the QaC data and find a similar pattern, though the results need caution due to a weak first stage.

I discovered other issues with PvW's sample in their code. Specifically, the estimation sample is conditional on educational outcomes, which is not innocuous. For example, they exclude observations with missing values on an alternative educational measure that they actually don't use in the final analysis. After including these observations in S12, the IV coefficient turns larger and significant, which suggests a non-randomness in the missings. A related problem is how PvW deal with implausible values in the main schooling measure, which is calculated as graduation year minus year of birth minus six. This first gives values between -57 and 62 and yields a weak first stage. PvW clean the lowest and highest values by comparisons with an alternative schooling measure that uses the instrument to determine schooling duration. Eventually, they exclude observations for whom the difference between the two variables is beyond the range from -2 to 4 years, which seems ambiguous and substantially reduces the sample. In S13, I apply an alternative data-cleaning procedure by dropping only the observations below the first and above the 99th percentile of the main schooling measure. This leaves individuals with

[9] For details, see Pischke (2007) and Online Appendix B. For simplicity, my $SSYs$ indicator does not distinguish between one or two years of exposure because the results hardly change. Table 1 does not report the results for the $SSYs$ variable but its effects on wages and schooling duration are negative as we might expect.

[10] This seems to be true also in the MC data as documented in Pischke and von Wachter (2005).

six to 21 years of schooling in the sample. This approach reduces the sample more sparingly and does not use the instrument for sample selection. It also generates significant wage returns of 8.4%. This result is not driven by the relatively low first stage because various alternative procedures yield similar IV estimates despite a stronger first stage.[11]

Overall, the results in Table 1 render the conclusion that compulsory schooling in Germany yields no wage returns at a minimum controversial. I also repeated the analysis for a subsample of individuals who eventually obtained at most the basic school degree. Similar to PvW, I found smaller and mostly insignificant IV estimates. However, PvW emphasize that this subsample yields unbiased estimates only if the reform did not affect completed track itself. Indeed, they find an insignificant effect on the probability of completing the basic track. I probe this finding in Table C.4 in Online Appendix C. Many of these reduced-form estimates suggest that the $C9$ reform increased the fraction of students who completed the middle track. This seems to be driven by a shift away from the basic track. This is not surprising given that extended compulsory schooling reduced the extra time in school needed to obtain the middle degree from two to one year.[12] Given the endogenous changes in educational careers, conditioning the sample on the eventually obtained school degree would lead to bias.

3 Concluding remarks

One of a few studies that obtain zero wage returns to compulsory schooling is Pischke and von Wachter (2008) for Germany. The authors argue that German students acquire labor market-

[11] Alternatively, I cut the sample at the 2nd and 98th percentiles, which corresponds to 7 and 19 years of schooling. I also obtain similar IV estimates from samples restricted to 6-20, 7-20, 6-19, 6-18 years of schooling.

[12] To mitigate the concern that my results are driven by the generally increasing trends in education, I additionally estimated specifications that control for the number of middle and academic schools using aggregate data from Kamhöfer and Schmitz (2016). The results remained nearly identical.

relevant skills much earlier than students in other countries. The zero effect on wages has been confirmed on different data by Kamhöfer and Schmitz (2016), who also find no effect on cognitive skills using a similar empirical approach.

This paper returns to the original estimates by Pischke and von Wachter (2008) and re-examines them. Specifically, I show that their estimates are sensitive to minor modifications in sample restrictions and model specifications. A further reassessment suggests that PvW's findings might suffer from previously unconsidered institutional details. Generally, my results cast doubts on the conclusion of zero wage returns to compulsory schooling in Germany.

Unfortunately, I cannot test the argument of no value added in terms of skills on my data. I find, nevertheless, that the reform under study might have increased the probability of completing the middle school track as opposed to the lowest track. In Germany, secondary school credentials essentially determine future occupational careers. Thus, signaling seems a plausible, though not necessarily exclusive, mechanism behind the potentially non-zero returns. Indeed, most recently, Hampf (2019) finds improved numeracy skills among the treated individuals.

While most of my estimates are in line with other studies on European data[13], this paper does not provide an uncontested prove for positive wage returns to compulsory schooling in Germany. This would require a careful development of an improved specification that addresses several problems, which complicated inference in the original study. I leave this challenging task for future research. Nevertheless, this paper yields implications for other studies examining non-pecuniary effects of the same reform because potential income gains might be an important channel through which compulsory schooling affects various outcomes.

[13]To give some specific examples; Brunello et al. (2009) obtain estimates between 0.035 and 0.095 depending on the quantile of ability and income. The IV estimates in Bhuller et al. (2017) vary between 0.045 and 0.115 depending on sample restrictions and earnings definition. Monetary returns to extensions within tracking systems seem to be rather at the lower end of this range albeit significant (e.g., Chib and Jacobi, 2016; Fischer et al., 2019).

Table 1: PvW's main estimates: replication and sensitivity tests

	(1) IV coeff. on years in school	(2) First stage $C9$	(3) First stage F-stat.	(4) Red. form $C9$	(5) Observations
Panel A: Pischke and von Wachter (2008)					
Original results	0.058	0.190 ***	23.73[a]	0.010	54,126
	(0.038)	(0.039)		(0.008)	
Replication	0.059	0.191 ***	23.11	0.011	54,096
	(0.038)	(0.040)		(0.008)	
Panel B: Sensitivity tests					
S1: Birth cohorts 1931-1960	0.067 *	0.187 ***	21.73	0.013	53,230
	(0.039)	(0.040)		(0.008)	
S2: Birth cohorts 1945-1960	0.096 **	0.239 ***	18.76	0.023 **	32,647
	(0.040)	(0.055)		(0.010)	
S3: Temporary extensions as additional instrument	0.080 **	0.187 ***	21.97	0.010	54,096
	(0.036)	(0.040)		(0.008)	
S4: Instrument comprises both temp. and final extensions	0.103 **	0.127 ***	22.61	0.013 **	54,096
	(0.042)	(0.027)		(0.005)	
S5: Age 19-60	0.065 *	0.187 ***	22.02	0.012	53,381
	(0.039)	(0.040)		(0.008)	
S6: Age 19-55	0.087 **	0.183 ***	19.99	0.016 **	50,293
	(0.041)	(0.041)		(0.008)	
S7: Age 25-65	0.088 **	0.187 ***	20.23	0.016 **	50,680
	(0.041)	(0.041)		(0.008)	
S8: Excl. bottom and top earners (0.5% each year)	0.075 **	0.192 ***	23.18	0.014 **	53,818
	(0.034)	(0.040)		(0.007)	
S9: Control for $SSYs$	0.069 **	0.248 ***	32.91	0.017 **	54,096
	(0.033)	(0.043)		(0.008)	
S10: Control for squared trends	0.110 **	0.181 ***	14.08	0.020 **	54,096
	(0.046)	(0.048)		(0.008)	
S11: Men only	0.132 *	0.121 ***	6.73	0.016 *	34,562
	(0.075)	(0.047)		(0.009)	
S12: Incl. missings on yrs of education	0.070 *	0.195 ***	25.92	0.014 *	56,895
	(0.037)	(0.038)		(0.008)	
S13: Yrs in school from 6 to 21	0.084 *	0.144 ***	9.80	0.012 *	59,702
	(0.047)	(0.046)		(0.007)	

Notes: Initial sample restricted to (West-)German citizens born 1930-1960 and aged 19-65. The dependent variable is log hourly wage. $C9$ and $SSYs$ are indicators for exposure to a ninth compulsory school year and the short school years, respectively. All regressions include a constant, indicators for state, year of birth, survey year, gender, and control for age (quartic), 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. [a]Pischke and von Wachter (2008) do not report the F-statistic directly but it can be approximately calculated using their first-stage estimation results as $(0.190/0.039)^2$.
Source: QaC 1979-1998/9, own calculations.

References

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017). Life-cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics 35*(4), 993–1030.

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

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

Chib, S. and L. Jacobi (2016). Bayesian fuzzy regression discontinuity analysis and returns to compulsory schooling. *Journal of Applied Econometrics 31*(6), 1026–1047.

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

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

Fischer, M., M. Karlsson, T. Nilsson, and N. Schwarz (2019). The long-term effects of long terms–compulsory schooling reforms in sweden. *Journal of the European Economic Association (forthcoming)*.

Grenet, J. (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.

Hampf, F. (2019). The effect of compulsory schooling on skills: Evidence from a reform in Germany. Ifo Working Paper 313, 01-67, Munich.

Ho, D. E., Z. C. Ashwood, and C. Handan-Nader (2019). New evidence on information disclosure through restaurant hygiene grading. *American Economic Journal: Economic Policy 11*(4), 404–428.

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

Kamhöfer, D. A., H. Schmitz, and M. Westphal (2019). Heterogeneity in marginal non-monetary returns to higher education. *Journal of the European Economic Association 17*(1), 205–244.

Kemptner, D., H. Jürges, and S. Reinhold (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *Journal of Health Economics 30*(2), 340 – 354.

Kesternich, I., B. Siflinger, J. P. Smith, and J. 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.

Margaryan, S., A. Paul, and T. Siedler (2019). Does education affect attitudes towards immigration? Evidence from Germany. *Journal of Human Resources*, doi: 56.2.0318–9372R1.

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

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

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

Piopiunik, M. (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, J.-S. (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, J.-S. and T. von Wachter (2005). Zero returns to compulsory schooling in Germany: Evidence and interpretation. NBER Working Paper 11414, Cambridge, MA.

Pischke, J.-S. and T. von Wachter (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *Review of Economics and Statistics 90*(3), 592–598.

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

Are there no wage returns to compulsory schooling in Germany? A reassessment

– Appendix –

Kamila Cygan-Rehm*

(University of Erlangen-Nürnberg, CESifo, IZA, LASER)

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

A Data and empirical strategy

The Qualification and Career Survey (QaC) is a repeated cross-section of the German labor force, and Pischke and von Wachter (2008) (henceforth PvW) use the first four survey waves collected in 1979, 1985/6, 1991/2, and 1998/9 (BIBB and IAB, various years).[1] They include individuals born between 1930 and 1960 who were between 19 and 65 years old when interviewed. Table A.1 shows the age structure in the data. Furthermore, they focus on German citizens living in the West German states (except for Berlin). They also exclude individuals with specific school degrees that could have been obtained only in socialist East Germany (GDR).

In the empirical approach, PvW exploit the regional variation in the timing of the compulsory schooling extensions from eight to nine years within an IV framework. While formally not presented in the original paper, the main outcome equation is

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

where y corresponds to the log hourly wages of individual i from state s in year t. The outcome variable is calculated from the gross monthly earnings (in DM) and the actual number of hours worked per week. $Educ$ represents an individual's years of schooling. German data do not typically report the number of years of schooling because the educational system is structured by tracks rather than by highest attended grade. Thus, PvW impute the years spent in primary and secondary school by using the information on the year of graduation from secondary school, the year of birth, and the typical age of school enrollment (six). χ and δ are vectors of state and year-of-birth fixed effects (FE), respectively. X comprises a quartic in age, an indicator for gender, and three indicators for the survey year. The corresponding first-stage equation is

$$Educ_{ist} = \pi C9_{st} + \lambda_s + \theta_t + X'_{ist}\phi + \nu_{ist} \tag{2}$$

where λ captures state FE and θ year-of-birth FE. $C9$ represents a binary instrumental vari-

[1]More recently, the survey was collected in the years 2005/6 and 2011/2.

able that indicates whether an individual was required to attend nine instead of eight years of compulsory schooling. PvW construct this variable by using the information on an individual's year of birth, the current state of residence, and the timing of the reform in each state (as shown column 1 in Table B.1 in Online Appendix B). Given that geographical information is limited to the current state of residence, the instrument suffers from a measurement error due to regional mobility after schooling completion. PvW argue that this measurement error would be problematic only if it is systematically correlated with the reform and wages, which is rather unlikely.[2]

Generally, the instrument varies over time and across states. To mitigate the concern that other factors disproportionately affected states over time, PvW control for state-specific linear trends in year of birth, which are included in X throughout. These trends and the year-of-birth FE should also flexibly capture the generally increasing trend in educational attainment (see Figure B.1 in Online Appendix B), which might potentially introduce changes in the share of compliers over time. Finally, the standard errors are clustered to for potential correlation among individuals from the same state and birth cohort.

Panel A of Table C.1 in Online Appendix C simply reprints PvW's first-stage results obtained by estimating equation 2. Panel A of Table C.2 reprints PvW's main estimates for the wage regressions. They begin by estimating equation 1 using standard OLS regressions. The main IV estimate is in column 4. Panels B of these tables show my exact replication of PvW's results.

[2]I investigated this issue using the QaC survey collected in 2005/6, which reports both the current state of residence and the actual state of school attendance. On average, for approximately 18% of respondents born between 1930 and 1960, the state of schooling differed from the state of residence at the time of the interview. Nevertheless, I did not find any evidence that regional mobility is significantly correlated with the reform indicator. Moreover, extending PvW's original sample by the 2005/6 data and using the available information on state of schooling does not appreciably change the IV estimates of wage returns to schooling. This is in line with PvW's argument that measurement error due to regional mobility should not lead to bias.

Table A.1: Age structure in the data by birth cohort and survey year

Birth cohort	QaC 1979	QaC 1985/6	QaC 1991/2	QaC 1998/9
1930	49	56	62	
1931	48	55	61	
1932	47	54	60	
1933	46	53	59	
1934	45	52	58	65
1935	44	51	57	64
1936	43	50	56	63
1937	42	49	55	62
1938	41	48	54	61
1939	40	47	53	60
1940	39	46	52	59
1941	38	45	51	58
1942	37	44	50	57
1943	36	43	49	56
1944	35	42	48	55
1945	34	41	47	54
1946	33	40	46	53
1947	32	39	45	52
1948	31	38	44	51
1949	30	37	43	50
1950	29	36	42	49
1951	28	35	41	48
1952	27	34	40	47
1953	26	33	39	46
1954	25	32	38	45
1955	24	31	37	44
1956	23	30	36	43
1957	22	29	35	42
1958	21	28	34	41
1959	20	27	33	40
1960	19	26	32	39

Source: Own illustration.

B Institutional background

Given the existing description by PvW, this section outlines only the most relevant details and additional background work, which motivates some of my robustness tests in Table 1 in the paper.

While the responsibility for educational policies lies with the federal states, the states negotiate and establish framework agreements to ensure the comparability of school systems. The duration of compulsory schooling was a much-disputed and controversial topic in the (West-) German educational debate after World War II (e.g., Leschinsky and Roeder, 1980). The Nazi regime centralized the education system and required at least eight years of schooling. After the war, Hamburg and Schleswig-Holstein were the first states that introduced a compulsory ninth grade ($C9$), which had already been mandated there before the war. The final implementation in the remaining states was staggered. Column 1 of Table B.1 re-prints the information on first birth cohorts affected by the reform from Pischke and von Wachter (2005).

In the postwar period, a weak labor market, high youth unemployment, and a shortage of apprenticeship positions for school graduates 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 local level to temporary prolong school attendance to 8.5 or nine years. The exact content of these laws differed across states. Most of them were limited to specific birth cohorts (see column 2 of Table B.1). Several of the laws were explicitly conditional on an insufficient number of apprenticeships relative to school graduates at the end of particular school years. In this case, each municipality could individually decide on the law's enforcement depending on the local labor market situation. The original legislation texts suggest that for students, the temporarily extensions were mandatory if their municipality actually made use of the law. Most of the laws also explicitly regulated at the state-level that the extensions should deepen general education and improve students' career choices.

Unfortunately, there is no systematic information on the actual implementation and enforcement of the temporary extensions at a local level. Nevertheless, official statistics on the ninth grade attendance in basic track at the state level might shed some light on this issue. I summarize the available data Table B.1 below. The data are available starting from the school year 1952/53, so that they might only capture a potential exposure to temporary extensions for children born 1939 and later. Note that in some states, the temporary extensions already affected children born 1933 (see column 2). Nevertheless, using the available data, I computed a ratio of the number of students in their ninth year compared to the number of students in their eighth year but in previous school year (i.e., from the same birth cohort). The numbers are presented in column 3, and suggest that the temporary laws were hardly ever implemented in Bavaria. However, especially, in Rhineland-Palatinate and Baden-Wuerttemberg the temporary extensions were a notable phenomenon. For example, in Rhineland-Palatinate, up to 93% of basic track students from cohorts 1938-1942 attended a ninth school year due to the local temporary extensions. Note that the first birth cohort affected by the universal introduction of the compulsory ninth year in this state was 1953.

Due to the recovering labor market in the mid-1950s, the demand for apprentices increased, which virtually suspended the temporary schooling extensions. In the 1950s, the political discussion focused mostly on 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 in the starting date of the school year from spring to fall. The change in schedule was completed within two short school years ($SSYs$) that actually jointly lasted only 16 months (for details, see Pischke, 2007). Thus, individuals affected by both the $C9$ reform and $SSYs$ had to remain in school longer than pre-reform but not as long as students exposed solely to the $C9$ reform. Save Hamburg, Lower Saxony, and Bavaria, all school children attending school in April 1966 experienced compressed school years. Although the $SSYs$ shortened the instructional time in the attended grade, they did not affect the curriculum. Bavaria postponed the implementation of the grade $C9$ 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).[3]

Generally, the $C9$ reform directly affected the duration of schooling among students who otherwise would have left school after eight years, which applies directly to the lowest secondary school track (basic track). However, the reform might also have obliged the potential dropouts from the other two tracks (middle and academic track) to remain in school longer because of extended compulsory schooling requirements. Figure B.1 shows the shares of students in each secondary school track across birth cohorts.

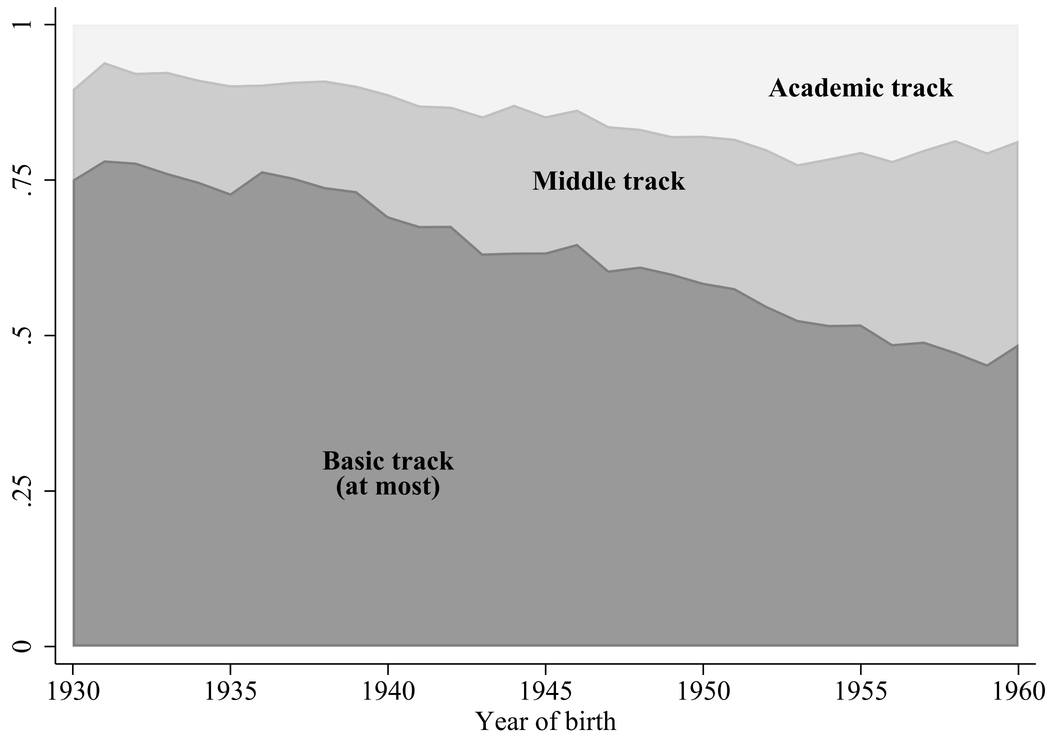
Table B.1: Final introduction of the ninth grade and earlier temporary extensions

	(1) First cohort with nine years (final reform)	(2) Birth cohorts affected by laws enabling temporary extensions	(3) Percentage of a cohort actually affected by temporary extensions
Schleswig-Holstein	1941		
Hamburg	1934		
Lower Saxony	1947	1935-1940	up to 15%
Bremen	1943		
North Rhine-Westphalia	1953	1935-1939	up to 20%
Hesse	1953	1933, 1936-1941	less than 5%
Rhineland-Palatinate	1953	1938-1942	up to 93%
Baden-Wuerttemberg	1953	1938-1941	up to 45%
Bavaria	1955	1938-1943	less than 1%
Saarland	1949		

Source: Column 1 from Pischke and von Wachter (2005). Column 2 based on original legislation from the state laws (further details available on request). Column 3 based on absolute numbers of basic track students in their eighth and ninth school year in school years 1952/53-1959/60 from *Statistisches Jahrbuch für die Bundesrepublik Deutschland* (various years). The digital version of the statistical yearbooks is available online at http://resolver.sub.uni-goettingen.de/purl?PPN514402342 [Last accessed: 25.09.2020]; own calculations.

[3] This might explain why I found some inconsistencies regarding the timing of the reform across various sources. However, this issue affects a limited number of cohorts from relatively small states and does not decisively drive PvW's findings. Thus, for comparability, I stick to their dates throughout.

Figure B.1: Shares of students by completed school track over time



Notes: This figure replicates Figure 1 from Pischke and von Wachter (2005) for birth cohorts 1930-1960.
Source: QaC 1979-1998/9, own calculations.

C Additional results

Table C.1: Replication: First-stage effect of the compulsory schooling reform on years in school

	Dependent Variable		
	Years in school (primary and secondary)		Attended basic track
	Full sample (1)	Basic track only (2)	Full sample (3)
Panel A: Pischke and von Wachter (2008)			
$C9$	0.190	0.285	-0.012
	(0.039)	(0.033)	(0.011)
Observations	54,126	32,970	54,126
Panel B: Replication			
$C9$	0.191	0.286	-0.012
	(0.040)	(0.033)	(0.011)
Observations	54,096	32,954	54,096

Notes: Sample restricted to (West-)German citizens born 1930-1960 and aged 19-65. Years in school are calculated as graduation year minus year of birth minus six. $C9$ is an indicator for exposure to a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, survey year, gender, and control for age (quartic). Robust standard errors clustered at state $\times$ year of birth cells in parentheses.
Sources: Panel A re-prints the top panel of Table 1 from Pischke and von Wachter (2008). Panel B based on own calculations using the QaC surveys 1979-1998/9.

Table C.2: Replication: Wage returns to schooling

	Full sample				Basic track only	
	OLS	OLS	Red. form	IV	Red. form	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pischke and von Wachter (2008)						
Years of education	0.061	-	-	-	-	-
	(0.001)					
Years in school	-	0.066	-	0.058		-0.005
		(0.002)		(0.038)		(0.034)
$C9$	-	-	0.010		-0.001	
			(0.008)		(0.010)	
Observations	54,126	54,126	54,126	54,126	32,970	32,970
Panel B: Replication						
Years of education	0.062	-	-	-	-	-
	(0.001)					
Years in school	-	0.066	-	0.059		-0.005
		(0.002)		(0.038)		(0.034)
$C9$	-	-	0.011		-0.002	
			(0.008)		(0.010)	
Observations	54,096	54,096	54,096	54,096	32,954	32,954

Notes: Sample restricted to (West-)German citizens born 1930-1960 and aged 19-65. The subsample of basic track students includes school dropouts. The dependent variable is log hourly wage. Years in school are calculated as graduation year minus year of birth minus six. Years of education (incl. postsecondary education and training) are constructed as the usual number of years spent in a particular educational route. $C9$ is an indicator for exposure to a ninth compulsory school year. All regressions include a constant, indicators for state, year of birth, survey year, gender, and control for age (quartic). Robust standard errors clustered at state $\times$ year of birth cells in parentheses.
Sources: Panel A re-prints the top panel of Table 2 from Pischke and von Wachter (2008). Panel B based on own calculations using the QaC surveys 1979-1998/9.

Table C.3: Replication: Sample means

	(1) Pischke and von Wachter (2005)	(2) Replication
Survey year	1986.9	1986.9
	(7.0)	(7.0)
Year of birth	1947.1	1947.1
	(8.5)	(8.5)
Age	40.1	40.1
	(10.0)	(10.0)
Female	0.36	0.36
Years of schooling (imputed from completed school track)	9.4	9.4
	(1.7)	(1.7)
Length of schooling (= year of graduation - year of birth - 6)	10.0	10.0
	(2.0)	(2.0)
Years of post-secondary training (imputed from completed vocational degree)	2.3	2.2
	(1.5)	(1.4)
Years of education (sum of imputed years of schooling and post-secondary training)	11.7	11.6
	(2.8)	(2.7)
Basic track: 8^{th} grade	0.40	0.40
Basic track: 9^{th} grade	0.21	0.21
Middle track: 10^{th} grade	0.23	0.23
Academic track: 12^{th} grade	0.03	0.03
Academic track: 13^{th} grade	0.13	0.13
Employed	1.00	1.00
Self-employed	-	0.07
Number of observations	54,126	54,096

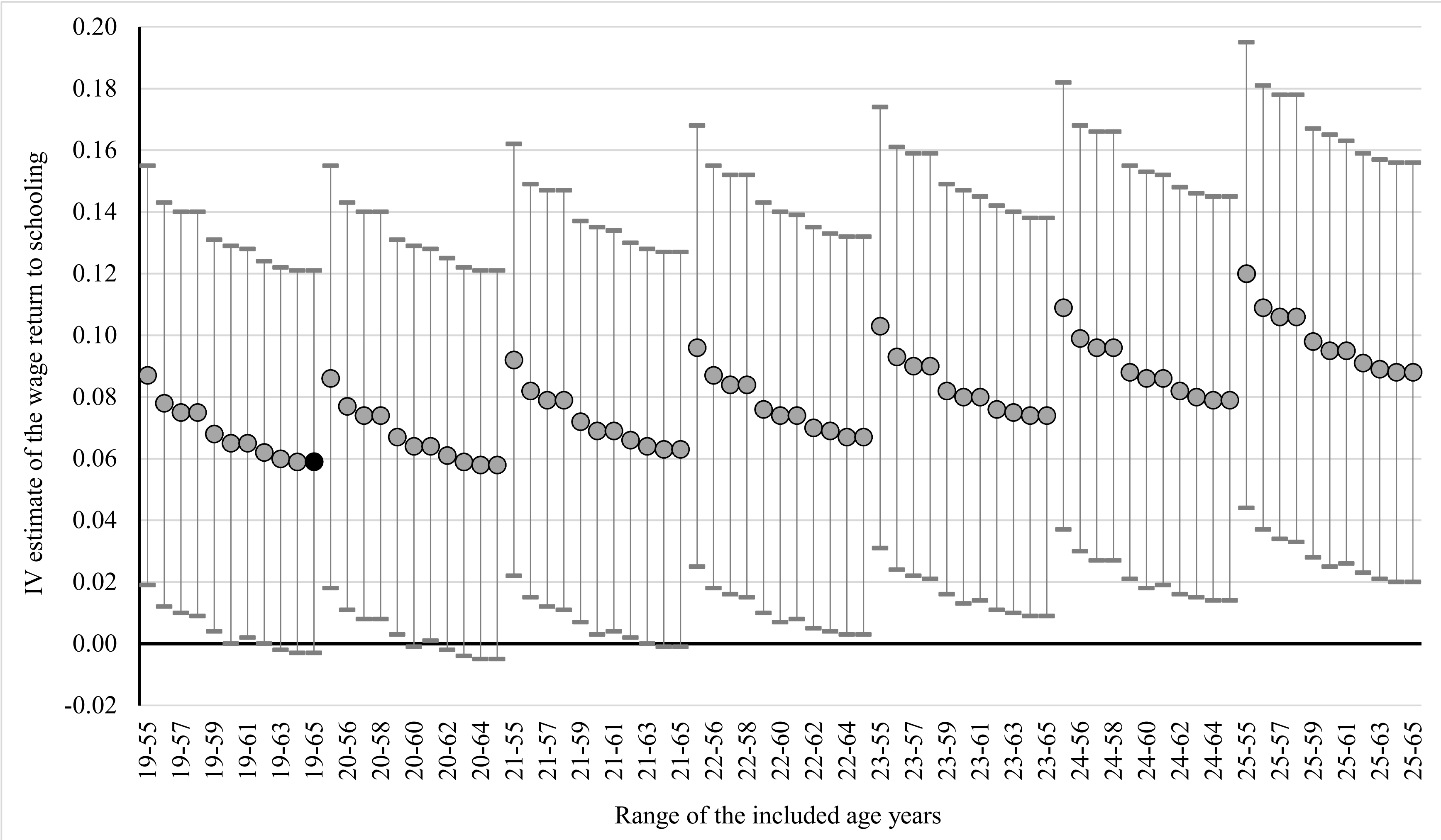
Notes: Sample restricted to (West-)German citizens born 1930-1960 aged 19-65. Track refers to completed school degree. Basic track includes school dropouts. Standard deviations in parentheses.
Source: Column 1 re-prints summary statistics from Appendix Table 1 in Pischke and von Wachter (2005). Column 2 based on own calculations using the QaC surveys 1979-1998/9.

Table C.4: PvW's reduced-form estimate of the effect of the ninth compulsory school year on completed school track: replication and sensitivity tests

	(1) Basic track (at most)	(2) Middle track	(3) Academic track	(4) Observations
Panel A: Pischke and von Wachter (2008)				
Original results	-0.012	n.a.	n.a.	54,126
	(0.011)			
Replication	-0.012	0.014	-0.002	54,096
	(0.011)	(0.009)	(0.008)	
Panel B: Sensitivity tests				
S1: Birth cohorts 1931-1960	-0.013	0.015 *	-0.002	53,230
	(0.011)	(0.009)	(0.008)	
S2: Birth cohorts 1945-1960	-0.029 *	0.026 *	0.004	32,647
	(0.015)	(0.013)	(0.010)	
S3: Temporary extensions as	-0.011	0.014	-0.003	54,096
additional instrument	(0.011)	(0.009)	(0.008)	
S4: Instrument comprises both	-0.015 **	0.012 **	0.002	54,096
temp. and final extensions	(0.007)	(0.006)	(0.004)	
S5: Age 19-60	-0.012	0.015	-0.003	53,381
	(0.011)	(0.009)	(0.008)	
S6: Age 19-55	-0.013	0.016 *	-0.003	50,293
	(0.011)	(0.009)	(0.008)	
S7: Age 25-65	-0.009	0.011	-0.002	50,680
	(0.012	(0.010)	(0.008)	
S8: Excl. bottom and top	-0.012	0.015 *	-0.003	53,818
earners (0.5% each year)	(0.011)	(0.009)	(0.008)	
S9: Control for $SSYs$	-0.008	0.011	-0.003	54,096
	(0.011)	(0.010)	(0.008)	
S10: Control for squared trends	-0.021	0.020 *	0.001	54,096
	(0.013)	(0.011)	(0.009)	
S11: Men only	-0.007	0.018 *	-0.011	34,562
	(0.012)	(0.011)	(0.010)	
S12: Incl. missings on years	-0.014	0.016 *	-0.002	56,895
of education	(0.010)	(0.009)	(0.007)	
S13: Years in school from 6 to 21	-0.010	0.016 *	-0.005	59,702
	(0.011)	(0.009)	(0.007)	

Notes: See Table 1 in the paper. The three outcome variables (in columns) correspond to indicators for mutually exclusive school tracks. Following Pischke and von Wachter (2008), basic track includes school dropouts. Source: QaC 1979-1998/9, own calculations.

Figure C.1: Effect of years of schooling on log wages across alternative ranges of included age years



Notes: Each dot shows the IV point estimate of the effect of years in school on log wages and the corresponding 90% confidence interval. The black dot highlights the age restrictions from Pischke and von Wachter (2008) (i.e., ages 19-65) . All regressions include a constant, indicators for state, year of birth, survey year, gender, and control for age (quartic), and state-specific linear trends in year of birth. Robust standard errors are clustered at state × year of birth cells. All samples are restricted to (West-)German citizens born 1930-1960.
Source: QaC 1979-1998/9, own calculations.

References

BIBB and IAB (various years). Bundesinstitut für Berufsbildung (BIBB), Berlin; Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesanstalt für Arbeit, Nürnberg: Qualifikation und Berufsverlauf 1979-1985/86. GESIS Datenarchiv, Köln. doi:10.4232/1.1243 (1979), doi:10.4232/1.12563 (1985/86), doi:10.4232/1.2565 (1991/92), doi:10.4232/1.12247 (1998/99)

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

Petzold, H.-J. (1981). Schulzeitverlängerung: Parkplatz oder Bildungschance? Die Funktion des 9. und 10. Schuljahres. Bensheim: Päd.-Extra-Buchverlag.

Pischke, J.-S. (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, J.-S. and T. von Wachter (2005). Zero returns to compulsory schooling in Germany: Evidence and interpretation. NBER Working Paper 11414, Cambridge, MA.

Pischke, J.-S. and T. von Wachter (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *Review of Economics and Statistics* 90(3), 592-598.