Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation

Jörg-Steffen Pischke, Till von Wachter, 2008

Radost Holler

Bonn Graduate School of Economics

December 20, 2018

Research Question

What are the (wage) returns to (an additional year of compulsory) schooling in West Germany?

Length of Compulsory Schooling and Reforms

- In the 1950's to 1960's the West German Bundesländer (states) increased compulsory schooling.
- The implementation date varied across Bundesländer.

Hauptschule:	**8**	$\rightarrow$	**9**
Realschule:	10	$\rightarrow$	10
Gymnasium:	13	$\rightarrow$	13

Potential Outcomes - Notation

Variable Definition:

Y_i	log wages
D_i	years of schooling
X_i	gender, age, age^2, year fe, year of birth fe, state fe, state fe $\times$ linear trend
Z_i	$\mathbb{1}\{$9 years of compulsory schooling$\}$

Potential Outcomes:

$$
\begin{aligned}
Y_{1i} &= \mu_1 + X_i\gamma + U_{1i} \\
Y_{0i} &= \mu_0 + X_i\gamma + U_{0i}
\end{aligned}
\tag{1}
$$

Here pretend years of schooling is a dummy.

Potential Outcomes

Observed Outcome:

$$
\begin{aligned}
Y_i &= D_i Y_{1i} + (1 - D_i) Y_{0i} \\
&= \alpha + X_i \gamma + \beta_i D_i + \varepsilon_i
\end{aligned}
\tag{2}
$$

where

$$
\begin{aligned}
\alpha &= \mu_0 \\
\beta_i &= \mu_1 - \mu_0 + (U_{1i} - U_{0i}) \\
\varepsilon_i &= U_{0i}
\end{aligned}
$$

Potential treatment state:

$$
D_i = D_{0i} Z_i + D_{1i} (1 - Z_i) \tag{3}
$$

where D_{ji} is the potential treatment state if $Z_i = j$

(Additional) Identification Assumptions, IV

A1 Independence: $\{Y_{1i}, Y_{0i}, D_{1i}, D_{0i} \perp\!\!\!\perp Z_i | X_i\}$

A2 Existence of a first stage:
$E(D_i|Z_i = 1, X_i) \neq E(D_i|Z_i = 0, X_i)$

A3 Monotonicity: $D_{1i} \geq D_{0i}\ \forall i$

Under these assumptions, we can identify the LATE.

This slide heavily relies on: Cornelissen et al. (2016)

(Additional) Identification Assumptions, IV

- A1 Independence: $\{Y_{1i}, Y_{0i}, D_{1i}, D_{0i} \perp\!\!\!\perp Z_i | X_i\}$
- A2 Existence of a first stage:
  $E(D_i|Z_i = 1, X_i) \neq E(D_i|Z_i = 0, X_i)$
- A3 Monotonicity: $D_{1i} \geq D_{0i}\ \forall i$

Under these assumptions, we can identify the LATE.

How do these assumptions relate to the assumptions made in the generalized Roy model? (for equivalence proof see Vytlacil, 2002)

This slide heavily relies on: Cornelissen et al. (2016)

The relationship between LATE and ATU

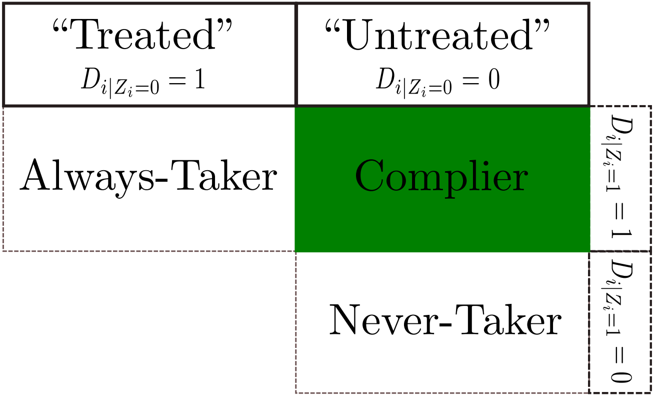
“Treated”	“Untreated”
$D_{i\|Z_i=0} = 1$	$D_{i\|Z_i=0} = 0$

The relationship between LATE and ATU

“Treated” $D_{i\mid Z_i=0} = 1$	“Untreated” $D_{i\mid Z_i=0} = 0$	
Always-Taker	Complier	$D_{i\mid Z_i=1} = 1$
Defier	Never-Taker	$D_{i\mid Z_i=1} = 0$

The relationship between LATE and ATU

+ Monotonicity assumption: no defier.



The relationship between LATE and ATU

- Monotonicity assumption: no defier.
- Perfect enforcement of reform: no never-taker

"Treated" $D_{i\|Z_i=0} = 1$	"Untreated" $D_{i\|Z_i=0} = 0$	
Always-Taker	Complier	$D_{i\|Z_i=1} = 1$

$$\Rightarrow \text{LATE} = \text{ATU} = (\text{PRTE})$$

Data (Constraints)

- Two data sets: Qualification and Career Survey (QaC) and Micro Census
- Main issue: neither data set includes years of education.

Imputation based on:

	QaC	Micro Census
year of birth	X	X
state of residence	X	X
year second. school graduation	X	
highest secondary degree	X	X
postsecondary education	detailed	not detailed

⇒ Only use imputed values for QaC.

The two-sample two-staged least square estimator (TSTSLS) - Intuition

- IV estimation of treatment effects relies on two moments/conditional means:
  1. GMM: Cov(Z, y),
     Wald-estimator: $E[Y_i|Z=1]-E[Y_i|Z=0]$
  2. GMM: Cov(Z, D),
     Wald-estimator: $E[D_i|Z=1]-E[D_i|Z=0]$

Zhao et al. (2017), Angrist & Krueger (1992), Inoue & Solon (2010)

The two-sample two-staged least square estimator (TSTSLS) - Intuition

- IV estimation of treatment effects relies on two moments/conditional means:
  1. GMM: Cov(Z, y),
     Wald-estimator: $E[Y_i|Z=1] - E[Y_i|Z=0]$
  2. GMM: Cov(Z, D),
     Wald-estimator: $E[D_i|Z=1] - E[D_i|Z=0]$
- Use the first stage of the QaC sample to estimate (2)
- and use the reduced form effect of the instrument on wages of the Micro Census to estimate (1)
- (here) the division of these two coefficients yields the TSTSLS estimator

Zhao et al. (2017), Angrist & Krueger (1992), Inoue & Solon (2010)

The two-sample two-staged least square estimator (TSTSLS)

A4 the data generating process of the samples is the same among the relevant dimensions (see Angrist & Krueger 1992, Zhao et al. 2017, for details)

A5 sample moments are independent

,

The two-sample two-staged least square estimator (TSTSLS)

A4 the data generating process of the samples is the same among the relevant dimensions (see Angrist & Krueger 1992, Zhao et al. 2017, for details)

A5 sample moments are independent

, The TSTSLS consistently estimates:

$$LATE^{MC,QaC} = \frac{E[Y_i^{MC}|Z_i^{MC}=1] - E[Y_i^{MC}|Z_i^{MC}=0]}{E[D_i^{QaC}|Z_i^{QaC}=1] - E[D_i^{QaC}|Z^{QaC}=0]} \tag{4}$$

$$\overset{A4}{=} \frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]} = LATE \tag{5}$$

Results

Independent Variable	Full Sample (1)
	Qualification and Career Survey
Dummy for cohort with ninth grade in basic track	0.190 (0.039)
Number of observations	54,126

	Dependent Variable: Log Wage			
	Full Sample			
Independent Variable	OLS (1)	OLS (2)	OLS (RF) (3)	IV (4)
	Micro Census			
Years of education	0.074 (0.001)	—	—	—
Imputed number of years in school	—	—	—	0.016 (0.015)
Dummy for cohort with ninth grade in basic track	—	—	0.003 (0.003)	—
Number of observations	939,736	939,736	939,736	939,736

Results

	Dependent Variable: Log Wage			
	Full Sample			
Independent Variable	OLS (1)	OLS (2)	OLS (RF) (3)	IV (4)
	Qualification and Career Survey			
Years of education	0.061 (0.001)	—	—	—
Imputed number of years in school	—	0.066 (0.002)	—	0.058 (0.038)
Dummy for cohort with ninth grade in basic track	—	—	−0.010 (0.008)	—
Number of observations	54,126	54,126	54,126	54,126

Discussion

- Results suggests that there are no wage returns to one additional year of schooling at the lower end of the schooling distribution in West Germany.
- Problem: according to my own research some of the reforms are incorrectly dated.

Discussion

State	Pivotal Cohort	
	Pischke & von Wachter	Laws
Schleswig-Holstein	1941	1940
Hamburg	1934	1934
Niedersachsen	1947	1947
Bremen	1943	1943
NRW	1953	1952
Hessen	1953	1947
Rheinland-Pfalz	1953	1951
BaWü	1953	1949
Bavaria	1955	1955
Saarland	1949	1952

References

Angrist, J. D. & Krueger, A. B. (1992), 'The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples', *Journal of the American Statistical Association* **87**(418), 328–336.

Angrist, J. D. & Pischke, J. S. (2009), Instrumental Variables in Action: Sometimes You Get What You Want, *in* J. D. Angrist & J. S. Pischke, eds, 'Mostly Harmless Econometrics - An Empricist's Companion', Princeton University Press, chapter 4, pp. 113 –218.

Cornelissen, T., Dustmann, C., Raute, A. & Schönberg, U. (2016), 'From LATE to MTE: Alternative methods for the evaluation of policy interventions', *Labour Economics* **41**, 47–60.
URL: *http://dx.doi.org/10.1016/j.labeco.2016.06.004*

Inoue, A. & Solon, G. (2010), 'Two-sample instrumental variables estimators', *Review of Economics and Statistics* **92**(3), 557–561.

Vytlacil, E. (2002), 'Independence, Monotonicity, and Latent Index Models: An Equivalence Result', *Econometrica* **70**(1), 331–341.

Zhao, Q., Wang, J., Bowden, J. & Small, D. S. (2017), Two-sample instrumental variable analyses using heterogeneous samples.
URL: *http://arxiv.org/abs/1709.00081*

Institutional Background

- Germany's secondary schooling system consists of three main tracks:
  - Hauptschule (basic track),
  - Realschule
  - Gymnasium (academic track).
- Children who graduate from Hauptschule or Realschule usually continue with an 3-year apprenticeship including part-time vocational education.
- Children who graduate from Gymnasium take part in a university entrance exam (Abitur)

Identified Parameters of IV with Covariates

- As discussed in the lecture, IV estimator without covariates yields the LATE. This is not generally true if we add additional covariates. (Cornelissen et al. 2016)
- If we partition the sample in subsamples s.t. for all $X_i = x$, then we obtain, the covariate specific LATE:

$$LATE(x) = E(Y_{1i} - Y_{0i} | D_{1i} > D_{0i}, X_i = x)$$

This slide heavily relies on: Cornelissen et al. (2016)

Identified Parameters of IV with Covariates

- As discussed in the lecture, IV estimator without covariates yields the LATE. This is not generally true if we add additional covariates. (Cornelissen et al. 2016)
- If we partition the sample in subsamples s.t. for all $X_i = x$, then we obtain, the covariate specific LATE:

$$LATE(x) = E(Y_{1i} - Y_{0i} | D_{1i} > D_{0i}, X_i = x)$$

- Applying the 2SLS with a **fully saturated 1st and 2nd stage** yields a **variance-weighted average of covariate specific LATEs.**
- Less saturated models seem to yield tolerable approximation (Angrist & Pischke 2009).

This slide heavily relies on: Cornelissen et al. (2016)