Replication materials for "Triangulating the relationship between education and attitudes towards immigration"
Journal of Experimental Political Science
Qinya Feng

The article uses individual level data from Swedish administrative registers. The specific register and survey sources used in the paper are: the Census data (Folk och Bostadsräkningarna, FoB) for 1960 and 1965, the LISA databases for 2009; the core cohort files and the SALTY survey from the Swedish Twin Register. The raw administrative register data is proprietary and cannot be shared. They can be obtained from Statistics Sweden and the Swedish Twin Register following the requisite ethics approval and application procedure. First, one needs to submit an application and obtain permission from a Swedish Ethical Review Board (https://etikprovningsmyndigheten.se/en/). Second, one needs to order data from Statistics Sweden and the Swedish Twin Register with at least one collaborator based at a Swedish university. More instructions of the ordering procedure can be found on the following sites. For Statistics Sweden, see https://www.scb.se/en/services/guidance-for-researchers-and-universities/. For the Swedish Twin Register, see https://ki.se/en/research/swedish-twin-registry-for-researchers.

In addition, the municipality data used in the Appendix analysis are sourced from the Election Data Archive (Valdataarkivet) for the years 1948-1970 and the Swedish Electoral Data (Valdata) for the years 1911-1944, obtained from the Swedish National Data Service (SND). The material in the Election Data Archive (1948-1970) was originally collected by Göran Gustavsson from the Department of Sociology at Lund University, while the material in the Swedish Electoral Data (1911-1944) was collected by Sten Berglund from the Department of Political Science at Umeå University. To prepare municipality-level aggregate results corresponding to the 1960 census version of municipality codes for merging with the municipality reform timing document (Holmlund, 2007), a workflow and processed dataset previously used in Lindgren et al. (2019) were adopted. The workflow processes the original data that mainly pertains to lower administrative units (primarily at the parish level), by merging with the official 1960 census municipality codes, addressing a series of changes in parish and municipality divisions and name changes over the period, and implementing other quality control measures.

All data processing and analyses were executed on a machine running Windows Server 2019 (Intel i9-7900X CPU @ 3.3GHz, 10 cores, 128GB RAM), on Stata/MP version 15.1, and R version 4.2.2.

The full replication package contains the following files:

File	Description
Data_EduIA.do	Loads raw data and prepares the main dataset used for all three studies. Produces EduIA.dta and Kommun.dta.

Analysis_EduIA.do	Runs all analyses in the main text. Produces Figure 2, Figure 3, Figure 4 in the main text. Also produces Table A.7 - A.8, Table A.13 - A.16, Table A.17 - A.20 in the Appendix.
analysis_logfile	The log file for running Analysis_EduIA.do.
Figure1.R	Produces Figure 1 in the main text.
Appendix.Rmd	Produces the Appendix.
Kommun.dta	Municipality level election data used in the Appendix.

Ntoe: The working directory "H:\Dofiles\Edu_IA" specified at the beginning of the files Data_EduIA.do, Analysis_EduIA.do, Figure1.R, and Appendix.Rmd was used by the author in the data construction and analysis processes. Directories to raw datasets appeared in these files are only accessible on the server where the raw administrative data is stored.

Description of the variables used in the empirical analyses, variable names correspond to the main dataset (EduIA.dta) used in *Analysis_EduIA.do* and *Appendix.Rmd*. See the Online Appendix for more detailed descriptions of the key variables.

LopNr: A unique personal identifier provided by Statistics Sweden that can be used to link individuals across different administrative registers.

LopNrParID: A unique twin pair identifier to identify twin pairs. From the core cohort file of STR.

Zyg: Zygosity of a twin pair. From the core cohort file of STR. Zyg = 1 for monozygotic twins, Zyg = 2 for dizygotic twins.

BirthYear: Year of birth. From the core cohort file of STR.

Age_2009: Age at 2009. Calculated based on year of birth.

female: female = 1 for female, female = 2 for male. From the core cohort file of STR.

old_BirthKommun: Birth municipality. Used the residence municipality from the census in 1960 (the earliest available).

sampleMZ: A sample indicator for Study 1 (complete MZ twin pairs).

refugee_r labor_immigration support_culture language_r: Four immigration related attitude items (unstandardised). Obtained from the SALTY survey from STR.

im_avg: An index for attitudes towards immigration based on the average of four items (standardised).

im_irt: An index for attitudes towards immigration based on the IRT model predicted score (standardised).

Education_year: Years of education. Obtained from the LISA database for the year 2009.

Tertiary_edu: Tertiary education. Tertiary_edu = 1 if years of education >= 15, Tertiary_edu = 0 if not.

Edu_orientation: Field of study for the higher level of education obtained. Obtained from the LISA database for the year 2009. 10 groups in total.

firstcohort60: A municipality's first birth cohort who were exposed to the education reform. Based on Holmlund (2007).

reformKommun60: Individual's residence municipality during the reform implementation period, based on the census in 1960. For individuals born in and after 1950, used the residence municipality from the census in 1965.

treatment: An indicator for reform treatment status. Constructed based on individual's residence municipality and reform time data collected by Holmlund (2007).

sampleReform: A sample indicator for Study 2.

PGI_EA: Polygenic index for education. Standardised by batch. Obtained from the polygenic index repository data from STR.

sampleDZ: A sample indicator for Study 3.

PC1-PC20: The first 20 principal components from the genetic relationship matrix. Obtained from the polygenic index repository data from STR.

LOC: Locus of controls based on a 12-item battery (standardised). Obtained from the SALTY survey from STR.

avg_RostBer, sd_RostBer: mean and standard deviation for municipality eligible voter population across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Valdelt, sd_Valdelt: mean and standard deviation for municipality turnout across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Andel_m, sd_Andel_m: mean and standard deviation for the Moderate party's vote share across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Andel_fp, sd_ Andel_fp: mean and standard deviation for the Liberal party's vote share across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Andel_c, sd_ Andel_c: mean and standard deviation for the Central party's vote share across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Andel_s, sd_ Andel_s: mean and standard deviation for the Social Democrats party's vote share across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

avg_Andel_v, sd_ Andel_v: mean and standard deviation for the Left party's vote share across 1940, 1944, 1948. Original data is from the Election Data Archive (Valdataarkivet) 1948-1970 and Swedish electoral data (Valdata) 1911-1944 obtained from the Swedish National Data Service (SND).

References

Berglund, S. (1988). Swedish electoral data 1911-1944 (1.0) [Data set]. Umeå University. Available at: https://doi.org/10.5878/000873

Gustafsson, G. (1988). Electoral data archive 1948-1970 (1.0) [Data set]. Lund University. Available at: https://doi.org/10.5878/000874

Holmlund, H. (2007). A researcher's guide to the Swedish compulsory school reform. *CEE Discussion Papers*.

Lindgren, K. O., Oskarsson, S., & Dawes, C. T. (2017). Can Political Inequalities Be Educated Away? Evidence from a Large-Scale Reform. *American Journal of Political Science*, 61(1), 222–236.

Lindgren, K.-O., Oskarsson, S., & Persson, M. (2019). Access to education and political candidacy: Lessons from school openings in Sweden. *Economics of Education Review*, 69, 138–148.