Appendix for "Triangulating the relationship between education and attitudes towards immigration"

Qinya Feng

Last compiled on 10 January, 2024

Contents

1 The SALTY survey, measurements, and descriptive statistics

During 2009 to 2010, the Swedish Twin Registry (STR) administered the SALTY survey (Screening across the Life-Span Twin Cohort Study) to twins born between 1943 and 1958. The SALTY was a collaboration between researchers from various disciplines, including epidemiology, medicine, and social sciences. In addition to questions relevant to multiple fields, the survey also covered numerous items on social and political attitudes and behaviors. A total of 24,916 twins were contacted with paper questionnaire for the survey, and 11,372 respondents provided informed consent to have their data stored and analyzed (Magnusson et al., 2013). Apart from immigration attitude items and the municipality reform timing document collected by Holmlund (2007), all other individual-level variables in the main analysis included in the main text are obtained from Swedish register data.

1.1 Measurements for education and immigration attitudes

This section provides more details on the measurements for education and immigration attitudes. Descriptions of other key variables (treatment indicator and PGI) are discussed in their respective sections.

1.1.1 Education

Information on education is obtained from LISA 2009 (Longitudinal integrated database for health insurance and labour market studies). Specifically, the "level" module in the Swedish education classification system (SUN 2000) records the highest level of education that one obtained as of 2009. Years of education is conversed based on the highest level of education. Years of education is operationalised based on the standard classification of education (ISCED) and in line with previous research with international comparison (Dinesen et al., 2016). Furthermore, the coding scheme is also consistent with some previous research

on the Swedish education reform where details on codings are publicly available (Hjalmarsson et al., 2015; Holmlund, 2007).

Specifically, the coding scheme is: (Old) compulsory school = 7 years; (New) compulsory school = 9 years; (Old) pre-senior secondary education = 9.5 years; High school = 10 or 11 or 12 years depending on program; Short university = 13 years; Longer university = 14 or 15 or 16 or 17 years depending on program, Short post-graduate degree = 18 years, Long post-graduate degree (PhD) = 20 years. In Study 1, it was mentioned that before the education reform, some schools already had 8 rather than 7 years of schooling. However, since school level information on their specific conditions at that time is not available, it is not feasible to distinguish individuals that had 8 years of schooling. Based on the Swedish education system, tertiary education corresponds to three years of post-secondary school education or longer, equivalent to 15 years of education.

1.1.2 Immigration attitudes

Immigration attitude indices are based on four items in the SALTY survey: 1) Increase labour immigration to Sweden; 2) Introduce a language test to become a Swedish citizen; 3) Accept fewer refugees in Sweden; 4) Increase the economic support to immigrants so that they can preserve their own culture. Respondents were asked whether they think these policy proposals are "very bad", "bad", "not good or bad", "good", "very good." We coded "very bad" as 1, and "very good" as 5. The responses to the language test and refugee acceptance question are reversely coded, so that higher values represent more open and supportive views on immigration policy arrangement.

The Cronbach's α for the four items is about 0.73, which indicates that the reliability/internal consistency of the items used is acceptable (Tavakol & Dennick, 2011). Item correlations are presented in Table A.1.

Table A.1: Item correlations

	Fewer refugees (r)	Increase labor immigraion	Language test (r)	Economic support
Fewer refugees (reversed)	1.00	0.52	0.51	0.41
Increase labor immigraion	0.52	1.00	0.28	0.37
Language test (reversed)	0.51	0.28	1.00	0.32
Economic support	0.41	0.37	0.32	1.00

[a] Notes: Fewer refugees (reversed) and Language test (reversed) are two items related to refugee acceptance and language tests that are reversely coded.

Figure A.1 shows the scree plot for the four items. It shows that there is a relatively large first factor or component, regardless of whether principal component or factor analysis is used. Factor loading shows that the variable on refugee acceptance has stronger correlation with the latent factor than other items.
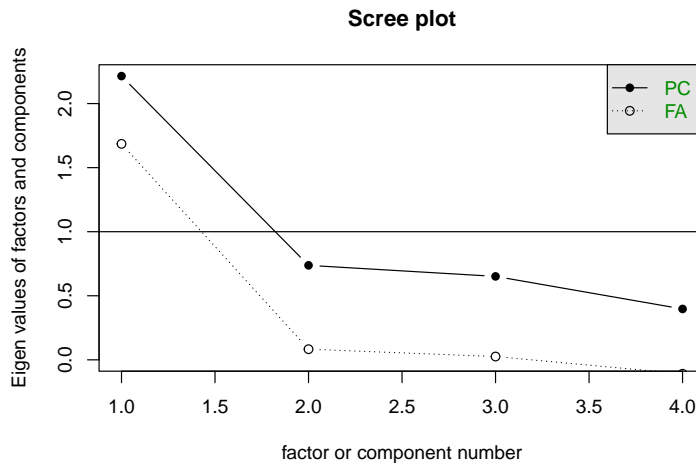
Figure A.1: Scree plot for immigration attitude items

Table A.2: Factor loadings

Variable	MR1
Fewer refugees (reversed)	0.848
Increase labor immigraion	0.602
Language test (reversed)	0.561
Economic support	0.538
SS loadings	1.686
Proportion Var	0.421

[a] Notes: Fewer refugees (reversed) and
Language test (reversed) are
reversely coded.

Apart from averaging responses to construct one index, I use the graded response model to estimate an latent variable based index. The graded response model is an extension of the 2 parameter IRT model to ordinal items, where items vary in their difficulty and discrimination. I use `irt grm` in STATA to fit the model, and use the default method empirical Bayesian mean to estimate the variable. Figure A.2 is the item characteristic curve for each item and response.

Figure A.2: ICC graph for immigration policy attitude items

Figure A.3 shows the distributions of both attitude indices in each sample. The general patterns are similar across samples.
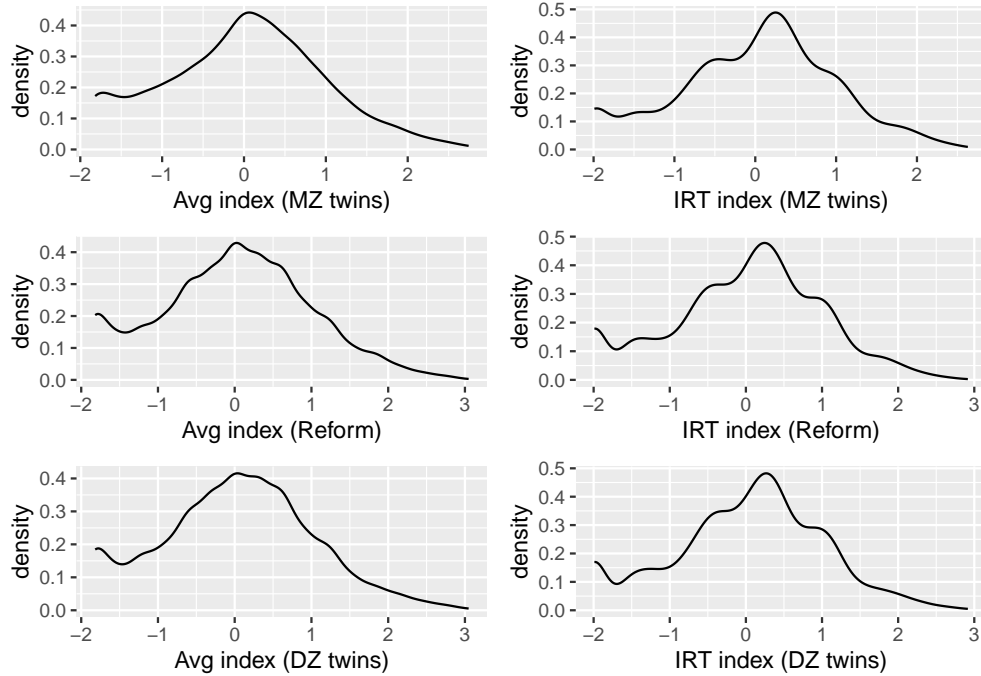
Figure A.3: Density plots for immigration attitudes in all samples.

1.2 Descriptive statistics

Table A.3 - A.5 give an overview of the samples and main variables. Observations with missing values on immigration attitude items are excluded. I use the STATA package REGHDFE to absorb fixed effects in all studies. Since REGHDFE drops singleton groups by default (Correia, 2017, 2023), the reform sample in Study 2 excludes those singleton observations that some birth year-municipality groups have. Twin samples in Study 1 (MZ twins) and 3 (DZ twins) are complete twin pairs.

Table A.3: Descriptive statistics for the MZ twin sample

	n	mean	sd	min	max
Immigration policy view (Avg)	1844	-0.02	0.99	-1.82	2.74
Immigration policy view (IRT)	1844	-0.01	0.99	-1.99	2.63
Fewer refugees (reversed)	1844	2.78	1.20	1.00	5.00
Increase labor immigraion	1844	2.64	1.00	1.00	5.00
Language test (reversed)	1844	2.33	1.19	1.00	5.00
Economic support for immigrant culture	1844	2.18	0.99	1.00	5.00
Years of Education	1844	12.15	2.58	7.00	20.00
Tertiary Education	1844	0.22	0.42	0.00	1.00
Female	1844	0.58	0.49	0.00	1.00
Age in 2009	1844	59.07	4.67	51.00	66.00

[a] Notes: The means and standard deviations for immigration attitude indices are slightly different from 0 and 1, as they are standardised based on responses from the whole SALTY sample.

Table A.4: Descriptive statistics for the reform sample

	n	mean	sd	min	max
Immigration policy view (Avg)	3755	-0.01	1.00	-1.82	3.05
Immigration policy view (IRT)	3755	-0.01	1.01	-1.99	2.92
Fewer refugees (reversed)	3755	2.77	1.21	1.00	5.00
Increase labor immigraion	3755	2.64	1.02	1.00	5.00
Language test (reversed)	3755	2.35	1.19	1.00	5.00
Economic support for immigrant culture	3755	2.19	1.00	1.00	5.00
Years of Education	3755	11.87	2.73	7.00	20.00
Tertiary Education	3755	0.20	0.40	0.00	1.00
Reform treatment	3755	0.41	0.49	0.00	1.00
Female	3755	0.54	0.50	0.00	1.00
Age in 2009	3755	60.39	3.69	54.00	66.00

[a] Notes: The means and standard deviations for immigration attitude indices are slightly different from 0 and 1, as they are standardised based on responses from the whole SALTY sample.

Table A.5: Descriptive statistics for the DZ twin sample

	n	mean	sd	min	max
Immigration policy view (Avg)	3168	0.02	1.00	-1.82	3.05
Immigration policy view (IRT)	3168	0.02	1.00	-1.99	2.92
Fewer refugees (reversed)	3168	2.79	1.21	1.00	5.00
Increase labor immigraion	3168	2.63	1.02	1.00	5.00
Language test (reversed)	3168	2.40	1.21	1.00	5.00
Economic support for immigrant culture	3168	2.22	0.99	1.00	5.00
Years of Education	3168	11.80	2.65	7.00	20.00
Tertiary Education	3168	0.18	0.38	0.00	1.00
EA PGI	3168	0.00	1.00	-3.45	3.53
Female	3168	0.54	0.50	0.00	1.00
Age in 2009	3168	59.17	4.54	51.00	66.00

[a] Notes: The means and standard deviations for immigration attitude indices are slightly different from 0 and 1, as they are standardised based on responses from the whole SALTY sample.

2 Study 1: Discordant MZ twins

2.1 Twin differences in educational attainment

Study 1 uses a sample of 922 MZ twin pairs. Table A.6 shows descriptive statistics of twin differences in education. Figure A.4 is the density plot of twin differences in education. Figure A.5 further shows the number of pairs for each twin differences in education.

Table A.6: Descriptive statistics MZ twin differences

	n	mean	sd	min	max
Immigration policy view (Avg)	922	0.75	0.61	0	3.65
Immigration policy view (IRT)	922	0.71	0.61	0	3.83
Years of Education	922	1.25	1.59	0	13.00
Tertiary Education	922	0.16	0.36	0	1.00
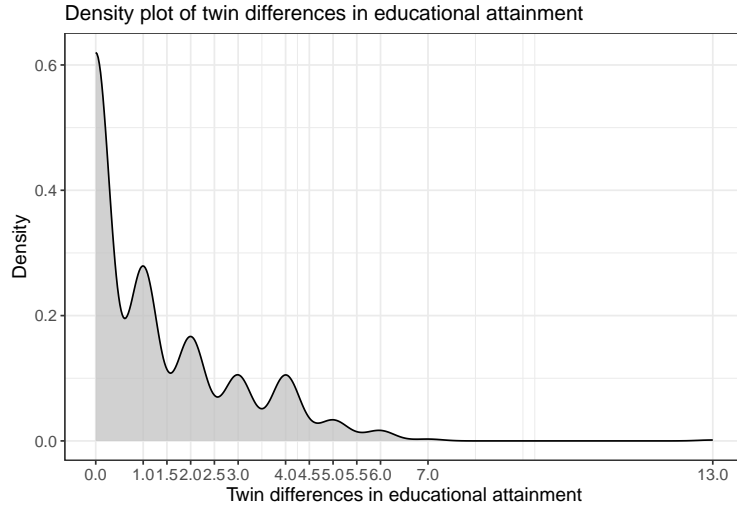
Figure A.4: Density plot of twin differences in educational attainment
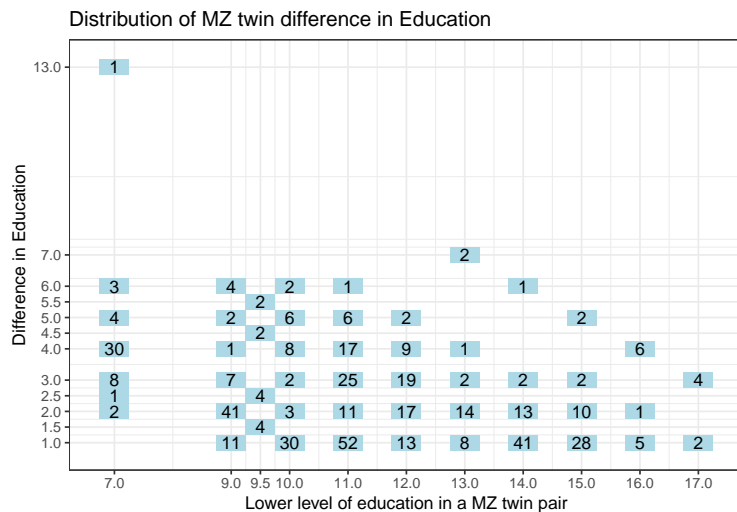
Figure A.5: Distribution for MZ twin differences in educational attainment

2.2 Sensitivity power analysis

Sensitivity power analysis is conducted under two sample sizes. One is the effective sample size based on the corresponding approximate intraclass correlation (ICC). The other one equals to the half of the total sample size. The results are presented in Figure A.6.

For the MZ twin sample (total N = 1844), the effective sample size is chosen as 1084, which is calculated based on an ICC of 0.7 (the ICC for years of education in the MZ twin

sample). The halved sample size is 922. I use an alpha level of 0.05 with two-sided t tests to plot the relationship between varying effect sizes and statistical power. The results show that, even under the halved sample size, there is over 85% power to detect an effect size of 0.1 (beta). Therefore, the MZ twin study should have decent power to detect a "small" effect size under the conventional standard. The beta coefficients for years of education obtained in the study are about 0.11-0.12.
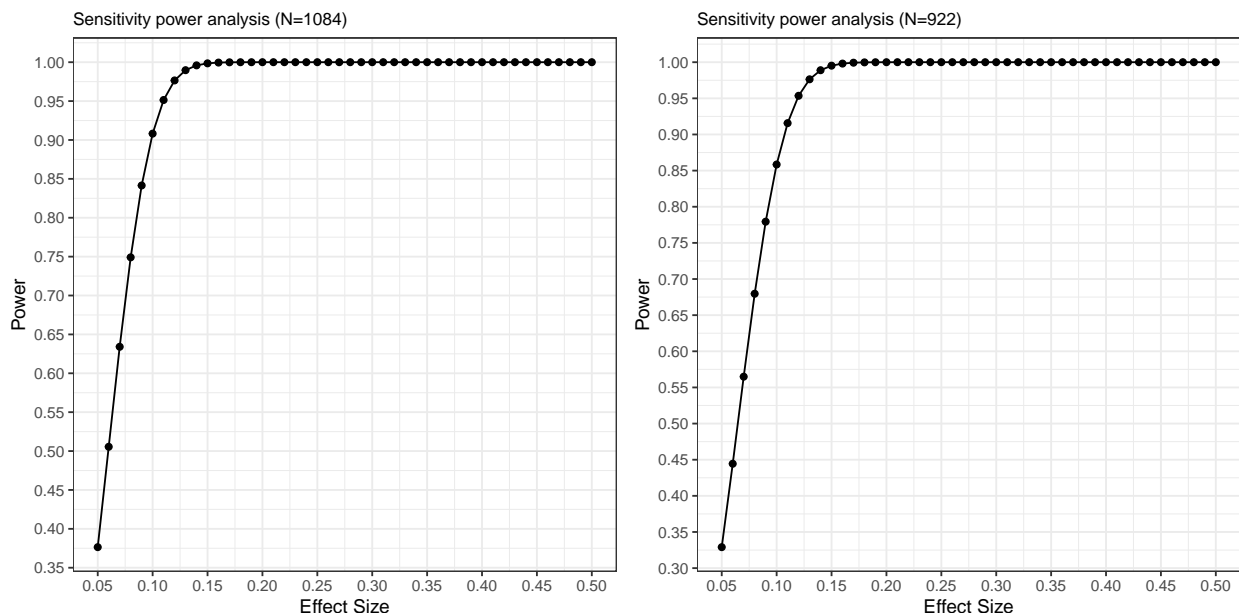
Figure A.6: Sensitivity power analysis for the MZ twin Study

2.3 Considerations on control variables

I include fields of study (highest degree) as an individual-level control, which has been over-looked in previous studies. Different fields of education have distinct values, structures, and resources that can influence attitudes towards immigrants differently. Stubager (2008) and Surridge (2016) argue that people educated in fields with a focus on the welfare of humans, such as arts and humanities, are likely to be more liberal than people who studied in fields focusing on managing people and manipulating objects, such as production and administration. Based on the social dominance theory, Sidanius et al. (2003) primarily dis-

tinguish between college majors that imply sympathy toward lower-status and subordinate social groups, and majors that inherit sympathy and assistance toward groups with power such as business executives. Apart from economic and technical skills, fields that prepare students with interactive and communicative skills may lead them to be exposed to others' views, thereby socialising them to exchange and tolerate divergent standpoints (e.g., Van de Werfhorst & De Graaf (2004)). Therefore, I expect that fields of study can vary in terms of their influence on the development of immigration attitude. Furthermore, individuals studying different fields may have varying considerations for their desired educational qualifications. Therefore, education orientation could potentially confound the relationship between educational attainment and immigration attitudes.

While there may be individual specific confounding factors, other suitable pre-treatment variables are not available or not sufficient in the cross-sectional survey data or registry data. Including variables such as cognitive performance in conscription records or grade variables at the 9th grade would significantly reduce the sample size, while conditioning on post-treatment factors can create spurious relationships and the bias on estimation can be in either direction (e.g., Acharya et al. (2016)).

Education orientation at one's highest level of education, as categorical variable, is identified by the "orientation" module in SUN 2000 (obtained from LISA2009, same as the educational attainment variable). Based on the official division (first digits in the coding scheme), fields of study can be divided into 10 groups: General education; Education science and teacher training; Humanities and art; Social sciences, law and business administration; Natural science, mathematics and Information and Communication Technologies (ICTs); Engineering, manufacturing and construction; Agriculture and forestry, veterinary; Health and welfare; Services; Unknown.

2.4 Full regression results

The effects of educational attainment on attitudes towards immigration (corresponds to Figure 2 in the main text)

Table A.7 presents results on the effects of education from within-family models and between family models. In the main text, the models include education orientation as a control variable; results from models excluding the variable are reported here.

In general, the education orientation variable does not change the results qualitatively. The estimates for education moderately fluctuate, and the within-family estimates for both years of schooling and tertiary education are higher in models without education orientation. We do see that several fields of study (compared with the reference group, General education) are significantly associated with immigration attitudes. However, only "Pedagogy and teaching" remain being significant in within family models.

Comparison with existing research

The effect size of interest for the MZ twin study can be considered in relation to the effect sizes of compulsory education reforms from existing research. The effects of one additional year of schooling on the attitude scale in the MZ twin study is about 0.04-0.05 standard deviation, which seems to be less considerable than the (pooled) reform effects. For example, using fuzzy RDD and pooling compulsory education reforms across 6 countries, Cavaille and Marshall (2019) found that one additional year of schooling decrease the anti-immigration attitude scales by about 0.2 standard deviations (they did not find significant effects for the Swedish reform, but noted that single country study could be underpowered). The authors interpreted the effect sizes as large, and suggested that the relatively uneducated reform-compliers (in the RDD) could be particularly susceptible to education's tolerance-inducing effects (p.261). D'Hombres and Nunziata (2016) also observed larger 2SLS estimates than the reduced form estimates. Similarly, they interpreted that there might be large effect of education among compliers - those with relatively low educational levels who are affected most positively by compulsory education reforms. They noted that the 2SLS estimates in

their study may be capturing a local average treatment effect (LATE) that is larger than the average treatment effect (ATE) (p.212). In other words, one year of education could be more consequential for the subgroup of individuals who would not have pursued further education otherwise, than for individuals who would have obtained university education anyways.

However, the seemingly small effect size in the MZ twin study may not be necessarily against the theoretical expectation that higher levels of education would be more effective in promoting pro-immigration attitudes. The wide range of within-twin-pair variation also implies heterogeneity, compared with the local variation (compliers) leveraged in RDD or IV analysis. Therefore, other than potential influences from sample and country differences, the small effects observed in the twin study perhaps also to some extent speak to the existing interpretation that LATE is more sizable.

Table A.7: The effects of education on immigration attitudes (MZ twins)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Avg index								
Years of Education	0.140***	0.051*	0.132***	0.056***				
	(0.015)	(0.020)	(0.012)	(0.016)				
Tertiary Education					0.463***	0.097	0.591***	0.136
					(0.082)	(0.087)	(0.076)	(0.083)
Pedagogy and teaching	-0.012	0.138			0.452***	0.303*		
	(0.137)	(0.148)			(0.128)	(0.127)		
Humanities and art	0.170	-0.142			0.589*	-0.008		
	(0.227)	(0.197)			(0.238)	(0.203)		
Social sciences Law Business Administration	-0.093	-0.069			0.243*	0.048		
	(0.096)	(0.102)			(0.095)	(0.091)		
Science Mathematics Data	-0.384	-0.002			0.112	0.178		
	(0.205)	(0.222)			(0.196)	(0.198)		
Technology and Manufacturing	-0.309**	0.072			0.004	0.176		
	(0.103)	(0.107)			(0.103)	(0.101)		
Agriculture Forestry Veterinary	-0.204	0.095			0.218	0.244		
	(0.241)	(0.256)			(0.214)	(0.241)		
Health care and Social care	-0.227*	-0.012			0.126	0.111		
	(0.105)	(0.122)			(0.098)	(0.106)		
Services	-0.139	0.042			0.175	0.169		
	(0.125)	(0.132)			(0.121)	(0.120)		
Unknown	0.196	0.081			0.496	0.200		
	(0.249)	(0.335)			(0.266)	(0.338)		
Constant	-1.562***	-0.655**	-1.670***	-0.698***	-0.206	-0.152**	-0.162	-0.046*
	(0.289)	(0.210)	(0.298)	(0.191)	(0.283)	(0.059)	(0.263)	(0.018)
R-squared	0.441	0.765	0.431	0.764	0.416	0.764	0.403	0.762
IRT index								
Years of Education	0.153***	0.049*	0.141***	0.054***				
	(0.015)	(0.020)	(0.012)	(0.015)				
Pedagogy and teaching	-0.079	0.125			0.432***	0.281*		
	(0.137)	(0.144)			(0.127)	(0.122)		
Humanities and art	0.100	-0.101			0.560*	0.026		
	(0.209)	(0.181)			(0.221)	(0.182)		
Social sciences Law Business Administration	-0.092	-0.059			0.276**	0.053		
	(0.098)	(0.098)			(0.097)	(0.085)		
Science Mathematics Data	-0.497*	-0.070			0.047	0.101		
	(0.216)	(0.209)			(0.214)	(0.188)		
Technology and Manufacturing	-0.300**	0.048			0.043	0.147		
	(0.106)	(0.105)			(0.106)	(0.102)		
Agriculture Forestry Veterinary	-0.179	0.123			0.284	0.266		
	(0.252)	(0.284)			(0.228)	(0.272)		
Health care and Social care	-0.218*	0.038			0.170	0.155		
	(0.108)	(0.120)			(0.099)	(0.103)		
Services	-0.129	0.097			0.215	0.218		
	(0.125)	(0.126)			(0.122)	(0.115)		
Unknown	0.118	0.045			0.447	0.158		
	(0.245)	(0.328)			(0.266)	(0.331)		
Tertiary Education					0.504***	0.095	0.620***	0.137
					(0.081)	(0.080)	(0.074)	(0.076)
Constant	-1.722***	-0.626**	-1.777***	-0.669***	-0.236	-0.148*	-0.163	-0.038*
	(0.291)	(0.202)	(0.294)	(0.181)	(0.285)	(0.058)	(0.260)	(0.017)
R-squared	0.452	0.780	0.444	0.779	0.423	0.779	0.411	0.777
Observations	1,844	1,844	1,844	1,844	1,844	1,844	1,844	1,844
Sex##Age	YES	/	YES	/	YES	/	YES	/
Orientation	YES	YES	NO	NO	YES	YES	NO	NO
Twin FE	NO	YES	NO	YES	NO	YES	NO	YES

16

Notes: Standard errors are clustered at twin-pair level. Between-family models control on age dummies, sex, and their interaction terms, as well as birth municipality fixed effects. The reference group for education orientation is General education. *** p<0.001, ** p<0.01, * p<0.05.

The effects of education by immigration attitude items (MZ twins)

Table A.8: The effects of education by immigration attitude items
(MZ twins)

VARIABLES	(1) refugee	(2) refugee	(3) labor	(4) labor	(5) culture	(6) culture	(7) language	(8) language
Years of Education	0.065**		0.043		0.053*		0.008	
	(0.023)		(0.022)		(0.023)		(0.027)	
Tertiary Education		0.118		0.206*		-0.059		0.055
		(0.096)		(0.091)		(0.100)		(0.121)
Constant	2.007***	2.644***	2.102***	2.506***	1.616***	2.155***	2.097***	2.170***
	(0.246)	(0.077)	(0.226)	(0.067)	(0.244)	(0.073)	(0.283)	(0.087)
Observations	1,844	1,844	1,844	1,844	1,844	1,844	1,844	1,844
R-squared	0.764	0.763	0.661	0.661	0.660	0.659	0.651	0.651
Sex##Age	/	/	/	/	/	/	/	/
Orientation	YES	YES	YES	YES	YES	YES	YES	YES
Twin FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Results from within-family models are reported here. Dependent variables are on the original 1-5 scale. The responses to the language test and refugee acceptance question are reversely coded. Standard errors are clustered at twin-pair level. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

3   Study 2: The Swedish education reform

In addition to a basic description of the Swedish education reform in the main text, here I provide more information and explanations about adjustments made.

The Swedish education reform has so far been employed to identify the causal effects of the reform and education in outcomes including educational attainment, earnings (Meghir & Palme, 2005), mortality (Meghir et al., 2018), criminal conviction (Hjalmarsson et al., 2015), and political candidacy (Lindgren et al., 2017). In contrast to a null effect of comprehensive education reforms on educational attainment in some other European countries (Cavaille & Marshall, 2019), the Swedish education reform indeed increased years of schooling by about 0.2-0.7 years according to these research, under varied research questions, research designs, and sample restriction procedures. The estimates from this study also fall in this range.

Reform assignment and sample restrictions

To generate the reform treatment indicator, I rely on the municipality reform timing

document coded by Holmlund (2007). This coding, based on various official documentations and archives, is used to match with the residence municipality of each individual. Following the recommendations provided (p.42-43) and previous practices (Hjalmarsson et al., 2015; Lindgren et al., 2017), I exclude individuals in municipalities, or parishes within municipalities, for which the documentation does not clearly identify the adoption date of the reform. For example, the reform was gradually implemented within some municipalities and the exact timings for sub-regions (parishes) of these municipalities were therefore unclear. Some measurement error in the coding is expected, but compared with other coding methods, Holmlund (2007) found that the reliability ratio of this coding is high and suggests its high quality (p.45-50).

In addition, the oldest cohort in the SALTY sample were born in 1943 thus we are not able to include elder cohorts. In Holmlund (2007), the reform affected birth cohort range was in fact also identified as between 1943 and 1955 (p.8-9) to ensure the accuracy of actual residence municipality for individuals (earlier records were not accurate enough).

Some other adjustments on sample restrictions were made to further alleviate concerns about individual self selection into (or out of) the reform and conflated effects from other policy changes over time. First, in each municipality, I restrict the sample to only include individuals born within the range of 10 years before or after the first cohort affected by the reform. This window width is decided by balancing between obtaining an adequate sample size for estimations and sufficiently narrowing the window size to exclude potential biases from other policy changes. 10-year is used for obtaining the main reform sample for analysis. In Table A.14, results using smaller window widths (9, 8, 7, 6 years) are also reported. Second, individuals who were one year older than the first affected birth cohort in each municipality were excluded. Hjalmarsson et al. (2015) found that the reform had significant effects on individuals who were born one year before the reform start years, which could be due to measurement error in reform assignment where some municipalities that partially introduced the reform to one year older cohorts are included; or grade repeators

that were educated in the new system were coded as untreated, thus producing the positive effects on education in this specific birth cohort. Alternatively, these individuals might start schooling later than they were supposed to be (Lindgren et al., 2017).

3.1 Treatment distribution by birth year and reform cohort

This section presents some overview about the phased-in introduction of the education reform.

Table A.9 shows the distribution of treatment status in Study 2 by individual birth year.

Table A.9: Treatment assignment by birth year

BirthYear	No. of individuals	% treated	No. of treated individuals
1943	310	4.19	13
1944	286	7.34	21
1945	374	5.08	19
1946	335	6.87	23
1947	323	16.72	54
1948	272	30.88	84
1949	316	41.46	131
1950	312	52.56	164
1951	262	62.60	164
1952	223	76.23	170
1953	240	87.08	209
1954	214	100.00	214
1955	288	100.00	288

[a] Notes: No.of individuals shows the number of individuals by each birth year cohort. % treated is the percentage, while No. of treated individuals is the number of individuals that were exposed to the education reform in the corresponding birth cohort.

Table A.10 shows the distribution of reform treatment status in each municipality reform cohort. For municipalities that only started the reform later on, the proportions of treated individuals gradually decrease from mid reforming municipalities to late reforming municipalities.

Table A.10: Treatment assignment by municipality reform cohort

Reform cohort	No. of individuals	% treated	No. of treated	No. of municipalities
1943	83	100.00	83	8
1944	46	100.00	46	9
1945	27	100.00	27	7
1946	169	82.25	139	14
1947	267	75.66	202	41
1948	510	59.61	304	54
1949	323	49.23	159	58
1950	426	50.94	217	46
1951	185	37.84	70	32
1952	571	28.20	161	98
1953	476	17.65	84	79
1954	245	12.24	30	50
1955	427	7.49	32	100

[a] Notes: No.of individuals shows the number of individuals in each municipality reform cohort (i.e., the first affected cohort in these municipalities are the same). % treated is the percentage, while No. of treated is the number of individuals that were affected by the reform in the corresponding birth cohort. No. of municipalities is the number of municipalities in the corresponding reform cohort.

3.2 Reform contents and result interpretation

As mentioned in the main text, the Swedish education reform also had significant pedagogical content changes (Hjalmarsson et al., 2015; Holmlund, 2007). Notably, placement into differentiated academic and nonacademic tracks was postponed. Compared with systems with early tracking, less tracking could allow intergroup contacts among students from different social backgrounds and increase social trust (Österman, 2021), which can be considered as general positivity towards strangers and may contribute to more positive views on immigration (Herreros & Criado, 2009). In addition, English was introduced as a compulsory subject from 5th grade in the new comprehensive school system (Holmlund, 2007). Increased and diverse social interactions (e.g., inter-ethnic friendship) could be part of the package of education, while second language acquisition and exposure might also lead to exposure to more information and foreign culture as well as influence perceptions of intergroup

relationships depending on contexts. These changes with potentially meaningful roles could all be contributing factors to the total effects of the reform on immigration attitudes.[1]

However, the primary factor making any effect could be the increase in years of schooling, because only a minority of cohorts differ from the majority in terms of their exposure to English learning and delayed tracking. In terms of English learning, English was included as a compulsory subject in the curriculum and introduced in the 5th grade in the new comprehensive school. Before the reform, in general pupils went through grades 1 to 4 or 1 to 6 in a common school (folkskolan). In either 4th or 6th grade, more able students were selected (based on past performance) for the five or three/four year long junior-secondary school (realskolan) (Holmlund, 2007). At this time, English was already a part of the curriculum for realskola, but it was not a compulsory part of the folkskola curriculum. In the autumn of 1953, English was made compulsory in the old folksola (Hjalmarsson et al., 2015). Thus, all treated and untreated individuals born in 1946 or later studied English starting in the 5th grade. In other words, only three oldest birth cohorts (born during 1943-1945) in the present study were subject to different exposure to English learning.

In the new 9-year comprehensive school, pupils were kept together in common classes longer than in the earlier school system, i.e., delayed tracking. However, as a compromise between the opponents of early tracking and its advocates, tracking remained for the 9th grade until 1969. Pupils would follow either a vocational track, a general track, or a theoretical track preparing for upper-secondary school (Holmlund, 2007, p.3). Meanwhile, elective courses were available to students in the 7th and 8th grades, so that students could prepare themselves for the track that they desired to follow in grade 9 (Hjalmarsson et al., 2015). In 1969, the 9th grade tracking was abandoned in favour of a completely comprehensive system (Holmlund, 2007), which only allowed ability grouping of pupils in mathematics and English

---

[1]This is also one reason for why some studies also focus on the total effect of the education reform, rather than using the reform assignment as an instrument for years of schooling. The reform assignment could influence the outcome not only through years of schooling but also through the aforementioned changes (and these changes were not just the result of increased schooling years), thus violating the exclusion restriction assumption (see e.g., Meghir et al. (2018)).

(Hjalmarsson et al., 2015). Therefore, all but the youngest two birth cohorts in the present study (born between 1954-1955) were subject to either the early tracking that existed in the pre-reform period, or the intermediate system with parallel classes that allowed for tracking within the new comprehensive school. The process of postponing tracking is in contrast to the Finnish education reform in the 1970s where tracking was an integral part of the transformation of Finnish educational system (Pekkarinen et al., 2009).

In all, although I cannot completely rule out the possibility that factors other than years of schooling partially contribute to the total effects of the reform, their role should not be exaggerated either (Hjalmarsson et al., 2015; Lindgren et al., 2017). Moreover, even though other research designs such as RDD and IV can be used to identify the effects of increasing educational attainment, they identify local average treatment effects of educational attainment on reform compliers. Therefore, it is likely that the effects of the Swedish reform under the current research design could be more comparable with the effects under the discordant twin design, than designs such as RDD and IV.

3.3 Sociopolitical characteristics of municipalities across reform cohorts

The identification of the reform effects rests on the assumption that, conditional on co-variates, the reform timing was uncorrelated with unobservables that could also influence immigration attitudes. Therefore, a main concern is that municipalities were not randomly selected into each experimentation round. Although the selection of municipalities was balanced by the committee that assessed the reform, reform implementation contained voluntarism elements and certain criteria might weigh more in the decisions (see Holmlund (2007)), thus leaving the reform assignment conditional on certain municipality characteristics.

There are several lines of evidence (from research using the reform time coding from Holmlund (2007)) suggesting that this assumption should be valid. First, Hjalmarsson et al. (2015: p. 1305-1306) show that there was no alarming difference in pre-reform trends

of years of schooling across municipalities. With more detailed analysis, they also show that individual's parental background does not predict reform treatment or have significant effects on years of schooling under appropriate specifications. Lindgren et al. (2017: p.225-227) also checked pre-reform trends of more municipality-level socioeconomic characteristics, finding that the average magnitudes of temporal change within reform cohorts were generally comparable. More nuanced balance tests also support the conclusion (in the Appendix of Lindgren et al. (2017)). Second, below I present visualisations that depict yearly sociopolitical characteristics of the 569 municipalities that appeared in Study 2, including voter turnout and the vote share of major active parties back then.[2] These visualisations reveal some differences among reform cohorts, but the temporal trends were similar, across pre and post reform periods (i.e., the disparities between reform cohorts were relatively consistent). Third, although I do not have information on municipality-level trends of shares of immigrants, the proportion of immigrants in the Swedish population was very low at the time ($<=5\%$) and immigrants at that time were mainly from Germany, other Scandinavian and Baltic countries. In all, the reform timing can be considered as plausibly exogenous. Furthermore, any time-invariant differences between early and late reformer groups will be accounted for by the reform cohort fixed effects.

The following graphs show the trends of sociopolitical characteristics (1940-1964) of municipalities for each reform cohort based on the means of municipal characteristics for each reform cohort at each election year (1940, 1944, 1948, 1952, 1956, 1958, 1960, 1964). The variables include voter population, turnout, and share of votes for major parties from the

---

[2]The original election result data are 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 generate municipality-level aggregate results aligned with the 1960 census version of municipality identifiers/codes, as used in the municipality reform time records (Holmlund, 2007), a workflow previously used in Lindgren et al. (2019) was adopted. This workflow processes the original data, which mainly pertains to lower administrative units (primarily at the parish level). It includes merging with the official 1960 census municipality codes and addressing changes in parish and municipality divisions and name changes over the period as well as other quality control measures (also see Lindgren et al. (2017)).

historical election results. These municipality-level statistics and their temporal trends are considered to capture important general municipal conditions. The staggered implementation period started from 1949 until 1962. In general, there is no clear sign that these socioeconomic conditions were associated with reform cohorts. One pattern that can be more clearly identified is that early reform cohorts had larger vote share for the Social Democratic party and lower voter share for the Agrarian party, than mid and late reform cohorts. However, the differences were relatively consistent over time.

In section 3.7, I also show that main model results with reform cohorts and birth cohorts fixed effects do not have substantial changes when controlling on these municipality level variables (including these municipality-level variables in the models with municipality fixed effects is not applicable).
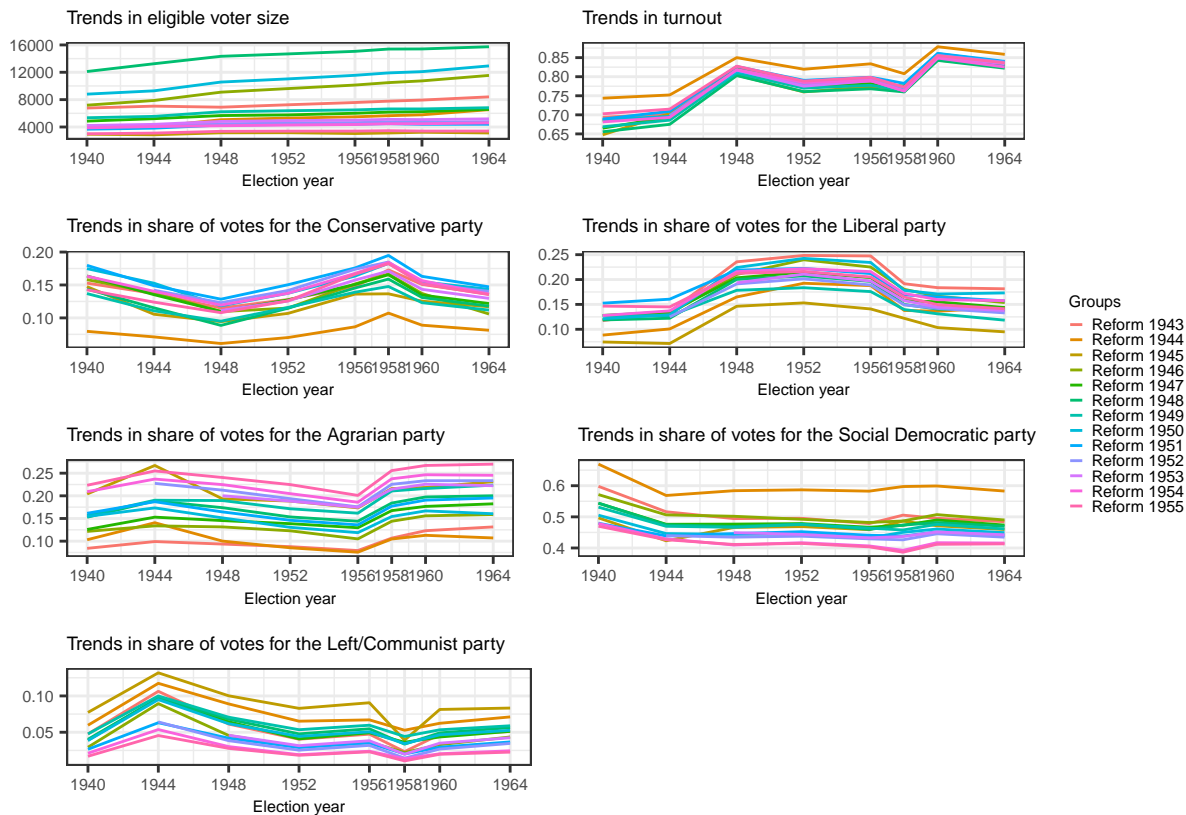
Figure A.7: Sociopolitical characteristics by reform cohorts

3.4 Descriptive statistics by reform treatment

This section provides descriptive statistics and Figure A.8 for educational attainments by respondents' reform treatment status as shown in the main text.

Table A.11: Descriptive statistics for the control group

	n	mean	sd	min	max
Reform treatment	2201	0.00	0.00	0.00	0.00
Years of Education	2201	11.54	2.89	7.00	20.00
Tertiary Education	2201	0.18	0.39	0.00	1.00
Immigration policy view (Avg)	2201	-0.06	0.98	-1.82	3.05
Immigration policy view (IRT)	2201	-0.06	0.99	-1.99	2.92
Female	2201	0.54	0.50	0.00	1.00
Age in 2009	2201	62.53	2.57	56.00	66.00

Table A.12: Descriptive statistics for the treatment group

	n	mean	sd	min	max
Reform treatment	1554	1.00	0.00	1.00	1.00
Years of Education	1554	12.35	2.41	7.00	20.00
Tertiary Education	1554	0.22	0.42	0.00	1.00
Immigration policy view (Avg)	1554	0.07	1.02	-1.82	3.05
Immigration policy view (IRT)	1554	0.07	1.02	-1.99	2.92
Female	1554	0.53	0.50	0.00	1.00
Age in 2009	1554	57.35	2.82	54.00	66.00
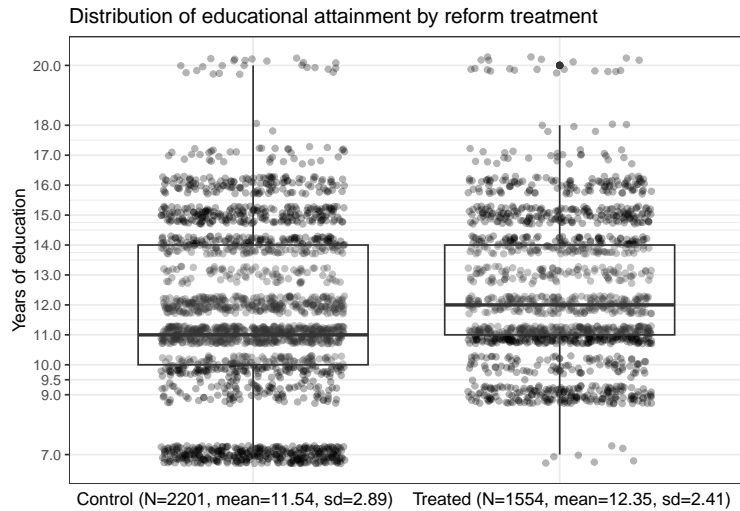
Figure A.8: Dsitribution of educational attainment by reform treatment

3.5 Reform effects on education (multiple window widths)

Similar as existing research that reported the reform effects on education, the estimates in this study suggest that the reform significantly increased years of schooling, although the magnitudes of effect size vary depending on window widths for sample restriction.

It is also necessary to note that the spill-over effects of the reform on obtaining higher education are likely to be limited. Previous population-wide analysis by Holmlund (2007) suggests a small spill-over effect of the reform beyond compulsory education, lasting for up to two years after the compulsory stage but not impacting tertiary education (p.13). Cavaille & Marshall (2019) also found no influence of education reforms on postsecondary education in their study (p.257).

Table A.13: Reform and years of education (mutiple window widths)

VARIABLES	(1) 10yr	(2) 10yr	(3) 10yr	(4) 9yr	(5) 9yr	(6) 9yr	(7) 8yr	(8) 8yr	(9) 8yr	(10) 7yr	(11) 7yr	(12) 7yr	(13) 6yr	(14) 6yr	(15) 6yr
Reform treatment	0.732***	0.461**	0.638***	0.733***	0.385*	0.536**	0.729***	0.298	0.491*	0.697***	0.370	0.632**	0.674***	0.196	0.394
	(0.134)	(0.166)	(0.188)	(0.135)	(0.170)	(0.192)	(0.135)	(0.179)	(0.206)	(0.134)	(0.190)	(0.216)	(0.134)	(0.206)	(0.239)
Constant	11.451***	11.560***	11.516***	11.484***	11.629***	11.587***	11.491***	11.678***	11.621***	11.565***	11.716***	11.622***	11.618***	11.849***	11.804***
	(0.109)	(0.117)	(0.103)	(0.112)	(0.121)	(0.107)	(0.116)	(0.125)	(0.113)	(0.117)	(0.132)	(0.122)	(0.121)	(0.138)	(0.131)
Observations	4,775	4,775	4,775	4,588	4,588	4,588	4,323	4,323	4,323	4,030	4,030	4,030	3,626	3,626	3,626
R-squared	0.029	0.035	0.270	0.030	0.035	0.269	0.029	0.036	0.273	0.026	0.031	0.274	0.027	0.033	0.287
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Cohort FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Reform Cohort FE	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO
Municipality FE	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES

Notes: Standard errors are clustered at munitipality level. Control variables include sex. *** p<0.001, ** p<0.01, * p<0.05.

3.6 Reform effects on immigration attitudes (multiple window widths)

Table A.14 presents the estimates for reform effects on immigration attitudes with multiple window widths for sample restriction. The estimates for reform treatment using a 10-year window shown in column 3 correspond to Figure 3 in the main text. As noted in the main text, the estimates for reform effects have large standard errors, and when narrowing the window width for sample restriction, the estimates from models controlling on reform cohorts or municipality fixed effects are not consistent in terms of effect direction. I consider it as mainly reflecting that the analysis is underpowered. If anything, given that a smaller window implies that the first affected cohorts get more weights, the decreasing estimates might suggest that the reform effects on the first few treated cohorts might be smaller than later treated cohorts.

Comparison with existing research

Two previous studies on the effects of education reform effects on immigration attitudes can be compared with Study 2 here. Cavaille & Marshall (2019) employed the same Swedish compulsory education reform and used a fuzzy RD design to exploit the discontinuity in years of schooling when the reform started to be implemented nation-widely in a later stage. Therefore, the sample was restricted to be young enough to be affected by the country-wide coverage around 1962 (earliest cohort born in 1951), and the study identified the local average treatment effects of additional schooling at the cutoff (N range from 1561 to 4759 depending on the outcomes). The estimates for LATE were not statistically significant, but the authors also noted that country-level analysis in their study might be underpowered. Finseraas et al. (2018) studied the effects of education reform in Norway where compulsory schooling was increased from 7 to 9 years and delayed tracking was also an element. Similar as the present study, they estimated the total effects of the reform within birth cohorts and municipalities. With a different composition of the attitude index and a slightly larger sample size (N = 4207), they found and argued that the reform had no causal impact on

immigration attitudes given substantively small point estimates and small standard errors.

Table A.14: Reform and immigration attitudes (multiple window widths)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
VARIABLES	10yr	10yr	10yr	9yr	9yr	9yr	8yr	8yr	8yr	7yr	7yr	7yr	6yr	6yr	6yr
Average index															
Reform treatment	0.143**	0.030	0.025	0.145**	-0.017	0.014	0.139**	0.003	0.047	0.138**	0.005	0.036	0.123*	-0.004	-0.033
	(0.049)	(0.069)	(0.080)	(0.049)	(0.073)	(0.085)	(0.049)	(0.079)	(0.094)	(0.049)	(0.088)	(0.106)	(0.050)	(0.093)	(0.112)
Constant	-0.079*	-0.030	-0.025	-0.075*	-0.003	-0.018	-0.076*	-0.013	-0.032	-0.068	-0.003	-0.015	-0.052	0.013	0.031
	(0.036)	(0.042)	(0.043)	(0.037)	(0.044)	(0.046)	(0.038)	(0.047)	(0.050)	(0.039)	(0.052)	(0.057)	(0.042)	(0.056)	(0.063)
Observations	3,755	3,755	3,755	3,598	3,598	3,598	3,382	3,382	3,382	3,155	3,155	3,155	2,852	2,852	2,852
R-squared	0.009	0.015	0.206	0.010	0.016	0.208	0.009	0.016	0.212	0.009	0.015	0.223	0.010	0.018	0.233
IRT index															
Reform treatment	0.136**	0.012	0.016	0.138**	-0.035	0.011	0.132**	-0.014	0.048	0.129*	-0.009	0.045	0.115*	-0.017	-0.012
	(0.050)	(0.071)	(0.083)	(0.050)	(0.074)	(0.088)	(0.050)	(0.081)	(0.097)	(0.050)	(0.090)	(0.108)	(0.051)	(0.094)	(0.116)
Constant	-0.075*	-0.021	-0.022	-0.070	0.007	-0.015	-0.070	-0.002	-0.031	-0.058	0.010	-0.014	-0.042	0.027	0.027
	(0.036)	(0.043)	(0.044)	(0.037)	(0.045)	(0.047)	(0.038)	(0.048)	(0.052)	(0.039)	(0.053)	(0.059)	(0.042)	(0.057)	(0.065)
Observations	3,755	3,755	3,755	3,598	3,598	3,598	3,382	3,382	3,382	3,155	3,155	3,155	2,852	2,852	2,852
R-squared	0.009	0.016	0.206	0.010	0.017	0.207	0.010	0.016	0.212	0.009	0.015	0.222	0.011	0.019	0.231
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Cohort FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Reform Cohort FE	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO
Municipality FE	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES

Notes: Standard errors are clustered at munitipality level. Control variables include sex. *** p<0.001, ** p<0.01, * p<0.05.

3.7   Reform effects with additional municipal controls

To evaluate whether pre-reform municipal characteristics affect the results, in Table A.15 I include the means and standard deviations of eligible voter population (RostBer, log transformed), turnout (Valdelt), and share of votes (Andel) for major parties (for the pre-reform period 1940-1948) as proxies for pre-reform municipal trends. For party vote share variables Andel, m denotes the conservative party, fp denotes the liberal party, c denotes the Agrarian party, s denotes the social democratic party, v denotes the left/communist party.

For immigration attitudes, only the turnout variable appear to be significantly related to immigration attitudes. Including it slightly decrease the estimates but does not substantially affect the results compared with Table A.14.

Table A.15: Reform effects with additional municipal controls

VARIABLES	(1) Avg index	(2) Avg index	(3) IRT index	(4) IRT index	(5) Education	(6) Education
Reform treatment	0.120*	0.022	0.112*	0.005	0.531***	0.441**
	(0.049)	(0.068)	(0.050)	(0.069)	(0.130)	(0.165)
(mean) RostBer	-0.001	0.013	-0.012	-0.007	0.094	0.094
	(0.042)	(0.044)	(0.043)	(0.043)	(0.131)	(0.118)
(mean) Valdelt	-1.009	-1.162*	-1.011*	-1.131*	-0.413	-0.386
	(0.516)	(0.515)	(0.514)	(0.514)	(1.423)	(1.462)
(mean) Andel_m	-5.115	-6.609	-5.372	-6.669	-16.914*	-17.744*
	(4.169)	(4.243)	(4.058)	(4.145)	(7.469)	(8.045)
(mean) Andel_fp	-5.233	-6.650	-5.554	-6.781	-18.085*	-18.601*
	(4.194)	(4.219)	(4.084)	(4.119)	(7.539)	(8.066)
(mean) Andel_c	-5.526	-6.835	-5.944	-7.092	-21.124**	-21.726**
	(4.191)	(4.223)	(4.078)	(4.120)	(7.475)	(8.018)
(mean) Andel_s	-5.501	-6.883	-5.910	-7.111	-19.167**	-19.802*
	(4.148)	(4.189)	(4.037)	(4.090)	(7.389)	(7.927)
(mean) Andel_v	-5.325	-6.581	-5.860	-6.904	-19.123*	-19.517*
	(4.286)	(4.296)	(4.158)	(4.187)	(7.654)	(8.232)
(sd) RostBer	0.670	0.780	0.628	0.724	2.135*	2.277*
	(0.411)	(0.410)	(0.396)	(0.389)	(0.941)	(0.976)
(sd) Valdelt	0.165	-0.031	0.097	-0.114	-0.784	-1.136
	(1.168)	(1.158)	(1.138)	(1.137)	(3.158)	(3.304)
(sd) Andel_m	1.363	1.387	1.123	1.034	-1.411	-1.400
	(1.628)	(1.543)	(1.624)	(1.539)	(5.337)	(5.389)
(sd) Andel_fp	-1.786	-1.283	-1.637	-1.113	-1.844	-1.599
	(0.966)	(0.967)	(0.966)	(0.955)	(2.918)	(3.066)
(sd) Andel_c	1.194	0.957	0.920	0.774	-3.009	-2.659
	(0.950)	(0.952)	(0.933)	(0.935)	(2.977)	(2.958)
(sd) Andel_s	-1.040	-1.065	-0.616	-0.625	0.092	-0.223
	(1.318)	(1.345)	(1.294)	(1.319)	(3.736)	(3.749)
(sd) Andel_v	2.372	2.110	2.277	1.874	-1.833	-2.783
	(1.842)	(1.834)	(1.780)	(1.757)	(4.796)	(4.834)
Constant	6.036	7.515	6.456	7.776	30.563***	31.248***
	(4.323)	(4.367)	(4.204)	(4.262)	(7.901)	(8.478)
Observations	3,755	3,755	3,755	3,755	4,775	4,775
R-squared	0.019	0.024	0.018	0.024	0.056	0.058
Controls	YES	YES	YES	YES	YES	YES
Birth Cohort FE	YES	YES	YES	YES	YES	YES
Reform Cohort FE	NO	YES	NO	YES	NO	YES
Municipality FE	NO	NO	NO	NO	NO	NO

Notes: The means and standard deviations of voter population (RostBer), turnout (Valdelt), and share of votes (Andel) for major parties (for the pre-reform period 1940-1948) are included in the model. Standard errors are clustered at munitipality level. Control variables include sex. *** p<0.001, ** p<0.01, * p<0.05.

3.8 Reform effects by attitude items

Table A.16: Reform effects by immigration attitude items

VARIABLES	(1) refugee	(2) refugee	(3) refugee	(4) labor	(5) labor	(6) labor	(7) culture	(8) culture	(9) culture	(10) language	(11) language	(12) language
Reform treatment	0.157**	0.033	0.043	0.107*	-0.018	-0.013	0.122**	0.102	0.096	0.085	-0.019	-0.045
	(0.061)	(0.086)	(0.102)	(0.051)	(0.070)	(0.081)	(0.047)	(0.065)	(0.073)	(0.055)	(0.080)	(0.092)
Constant	2.701***	2.756***	2.754***	2.583***	2.636***	2.634***	2.098***	2.108***	2.117***	2.335***	2.379***	2.388***
	(0.043)	(0.051)	(0.053)	(0.035)	(0.041)	(0.041)	(0.032)	(0.036)	(0.039)	(0.042)	(0.048)	(0.048)
Observations	3,755	3,755	3,755	3,755	3,755	3,755	3,755	3,755	3,755	3,755	3,755	3,755
R-squared	0.010	0.016	0.199	0.007	0.014	0.199	0.007	0.010	0.194	0.011	0.017	0.188
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Cohort FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Reform Cohort FE	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO
Municipality FE	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES

Notes: Standard errors are clustered at munitipality level. Dependent variables are on the original 1-5 scale. Control variables include sex. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

4 Study 3: DZ twins and EA PGI

In addition to a minimum explanation on PGI and the application of it in social research in the main text, here I provide a primer about GWAS, PGI, and the polygenic index repository. I also discuss sources of bias and measurement error in PGIs, the interpretations of PGIs, and some examples of using PGIs to investigate social mechanisms.

4.1 Polygenic index

Genome-wide association studies

Genome-wide association study (GWAS), emerged in 2005, has allowed researchers to identify associations between thousands and millions of genetic variants throughout the genome. This means that researchers are able to actually measure and analyse genetic variants linked with traits of interest across unrelated individuals. Genetic variants here are the variations on DNA sequence between individuals. And the most common form of genetic variants in humans is single nucleotide polymorphism (SNP), where individuals differ only by one nucleotide. The nucleotide of a SNP that is more common in a population is called

the major allele, and the less common one is called the minor allele. Humans receive two alleles, one from each parent. In this case, a person might have 0, 1, or 2 copies of minor allele. This number represents the genotype for a specific SNP. Genotyping thus means to identify these genetic variants that an individual possesses. Overall, humans are over 99% genetically identical in SNPs, only the remaining SNPs contains information about human variations.

In a conventional OLS-based GWAS, a specific trait of interest, such as height, is regressed onto each SNP with standard covariates like age, sex, and principal components computed from the genetic relationship matrix. This entails running millions of regression models to identify associations between SNPs and a trait. A genome-wide significant SNP needs to have a p-value (for its association with the target trait) that is lower than a multiple testing adjusted significance threshold, typically, $p < 5 \times 10^{-8}$. With developments in computing and statistical techniques and decreasing genotyping costs, GWAS research has expanded rapidly, with large sample sizes and a wide variety of traits such as educational attainment, cognitive ability, psychological traits, physical and health related traits. Large-scale GWASs often identify hundreds of or over a thousand of genetic variants that are significantly associated with the trait in the study, with each genetic variant showing very small effect on the trait - a phenomenon called polygenicity. For example, for the first three discovered SNPs in one of the first GWASs on educational attainment, each explained only about 0.02% of phenotypic variance in education, which corresponds to a difference of about one month of schooling per allele (Rietveld et al., 2013).

This theoretical and methodological development also explains why earlier studies on candidate genes often fail to be replicated. Candidate gene studies only focus on a handful of predefined genes that are hypothetically relevant to the trait of interest, such as the relationship between 5HTTLPR (a genetic variant of the serotonin transporter gene) and depression (Caspi et al., 2003). Since the influences from only a handful of genes are likely to be very small, candidate gene studies based on a few thousand individuals were dramatically

underpowered to identify genetic influences. The relationship identified in those studies can thus be false positive results.

PGI construction

PGI represents the aggregated genetic effects across the genome by summing the alleles that an individual carries and weighting them by the allelic effect size estimates from GWASs. As a simplified example for illustration, we can consider a case of only two SNPs in the genome. As aforementioned, for each SNP a given individual can have either 0, 1 or 2 minor alleles. If we assume that GWASs on educational attainment show that each additional minor allele for the first SNP is significantly associated with 1 week of increase in educational attainment, and 2 weeks of increase in educational attainment for the second SNP. Then, a PGI for education can be calculated by adding the number of minor alleles for the first SNP multiplied by 1, and the genotype for the second SNP multiplied by 2. The sum is therefore a simple measure of an individual's genetic predisposition towards the number of weeks of educational attainment.

Note, however, PGIs can be constructed in various ways although the essence of the idea remains the same. For example, PGI construction can vary in terms of the choice for GWAS discovery sample, criteria on which SNPs are included, and corrections for linkage disequilibrium (LD)[3] (see discussions in, for example, Muslimova et al. (2023)).

Current limitations of PGI

PGIs does not fully capture genetic variation related to target traits. The predictive power of PGI is not quite satisfactory if we compare heritability estimates from genotype data (the proportion of trait variation accounted for by measured genetic variation) with the heritability estimates from twin studies. While heritability based on all genotyped genetic variants (SNPs) has been shown to be around 20% (Rietveld et al., 2013), a meta-analysis of twin studies reported that about 40% of the variation in educational attainment can be attributed to genetics (Branigan et al., 2013).

---

[3]If an allele of one genetic variant is inherited or correlated with an allele of a nearby genetic variant within a given population, these alleles are in linkage disequilibrium.

The main explanations for the gap between two heritability estimates are that complex traits such as social and behavioural outcomes are 1) highly polygenic (i.e., are influenced by many genetic variants), and could be contributed by 2) rare variants and 3) non-additive variance; and another explanation is that twin studies overestimate heritability (see, e.g., (Young, 2019)). It is expected that, as the sample sizes in GWAS increase, more comprehensive common genetic variants can be identified and incorporated in PGI construction. Genotyping rare variants and identifying non-additive variance remain challenging in the investigation of genetic architectures. However, in terms of non-additive variance, recent research has argued that genetic variance for complex traits is predominantly additive (e.g., (Hivert et al., 2021)). Currently, the predictive power of PGIs for social and behavioural outcomes is generally lower than traits such as height and BMI. And it is acknowledged and cautioned that PGIs are poor predictors and *should not be used in predicting* **individual** *social and behavioural outcomes* (see e.g., (Becker et al., 2021)).

EA PGI

Advancements in GWASs offer promise for enhancing the precision of PGIs through improvements in discovery samples and methodologies. The iterations of the polygenic index for educational attainment is one example. The polygenic index has iterated from EA1 in 2013, EA2 in 2016, EA3 in 2018, to EA4 in 2022. EA1 had a sample size of about 100,000 individuals and the polygenic index predicted about 2% of the variance in years of schooling. The most recent GWAS meta-analysis for education, EA4, has a combined sample size of about 3 million individuals and identifies 3,952 approximately uncorrelated genome-wide-significant SNPs. The EA PGI constructed explains about 12-16% of variance in education (Okbay et al., 2022).

The EA PGI in the polygenic index repository was constructed from a GWAS of about 1 million individuals. The incremental R^2 of (single-trait) EA PGI range between 6% and 11% across five validation datasets in (Becker et al., 2021) (note that the same PGI may have different predictive power in different samples if there are differences in the phenotype

measure). In Dawes et al. (2021), the authors compared the incremental R^2 for EA PGI with other predictors of voter turnout in five large samples from the US and Sweden. Although the results vary across samples and different measures for turnout, the incremental R^2 for EA PGI in some cases are on par with variables like parental education and parental income which are often used as proxies for family socioeconomic status in social research.

4.2   The polygenic index repository

The polygenic index repository (Becker et al., 2021) contains PGIs for 47 phenotypes and is based on meta-analysis with data from 11 participating datasets (such as the UK biobank and other registries with individuals mainly of European ancestry). It has been established with a uniform methodology and made available to help overcome obstacles faced by researcher who want to conduct GWASs and use PGIs. For example, there are administrative hurdles in accessing genetic data and summary statistics, inconsistencies of methodologies used in PGI construction, overlap between GWAS discovery sample and the target dataset).

To access the PGIs and Principal components in a dataset, researchers will need to follow the data access procedures for the specific dataset. The up-to-date procedures can be found on the Social Science Genetic Association Consortium (SSGAC) website (https: //www.thessgac.org/ pgi-repository).

The STR (genotyped individuals) is one of the participating datasets.To avoid inflation of estimates due to sample overlap, the STR cohort was excluded in the GWAS discovery stage before constructing EA PGI for individuals in the STR cohort (i.e., out-of-sample prediction). Researchers interested in using STR data must obtain approval from the Swedish Ethical Review Authority and from the Steering Committee of the Swedish Twin Registry. Researchers using STR data are required to follow the terms of a number of clauses designed to ensure protection of privacy and compliance with relevant laws. For further information please visit ki.se/en/research/the-swedish-twin-registry.

To assist interested researchers to get familiar with and use the repository, the PGI repos-

itory also provides a user guide, with more specific accounts on methodologies (and software) for PGI and principal component construction; genotyping, imputation, and phenotype definitions in Repository datasets; a list of predictive power (incremental R^2) of Repository PGIs in validation datasets. The authors also provide a list of suggestions on how to think through some of the interpretational issues when using PGIs.

In addition to theoretical and technical explanations, a FAQ section about the repository's research background, design, social and ethical implications, in a more accessible form, is appended in the supplemental material of Becker et al. (2021).

4.3 Applications of PGI

Here I discuss several aspects regarding the applications and interpretations of PGI which are simplified in the main text. In the end, I also include some recent research adopting novel approaches to employ PGIs.

First, PGI is subject to measurement error and represents a noisy measure of (direct) genetic effects. The association between a genetic variant and a phenotype in most GWASs captures the effect of the variant, but likely also captures the indirect effect of the variant and confounding effects (Young et al., 2019). Indirect effects could be genetic nurture. For example, offspring genotypes can be correlated with parental genotypes (or other siblings/relatives' genotypes), while parental genotypes can influence parental behaviours which in turn exert influence on offsprings's phenotypes (Cawley et al., 2020; Wu et al., 2021). Genetic nurture therefore can indirectly cause association between a variant and a phenotype. Other confounding effects could be assortative mating and population stratification. Assortative mating occurs when partners that produce offspring select each other on the basis of particular trait (e.g., education, BMI), which may induce a fraction of the association between a genetic variant and a trait to be contributed by other genetic variants that are also correlated with the trait (inflate the effect size of the genetic variant). Population stratification occurs when there are regional differences/other ancestral differences in allele frequency

related to a trait of interest. If not adequately controlled for by principal components, residual population stratification can be correlated with both phenotypes and genotypes. Given these issues, family-based GWASs and other methodologies have been developed or conducted to gradually eliminate these sources of bias in GWASs (Howe et al., 2022; Young et al., 2018). In all, PGIs constructed based on effect sizes estimated in most GWASs captures some true genetic effects but also indirect genetic effects and other residual confounding.

Second, when applying PGIs to downstream analysis, the association between a PGI and a trait might be confounded by environmental mechanisms as discussed in the previous paragraph. But in within-family analysis, as noted, genetic inheritance from a pair of biological parents is random. Genetic differences between siblings therefore are uncorrelated with familial genetic influences that shape environments, or population stratification and assortative mating. Consequently, within-family analysis identifies the causal effects of a PGI. Since PGIs contain measurement error, within-family estimates of PGI as genetic effects are likely subject to attenuation bias (Becker et al., 2021; J. Fletcher et al., 2021; Trejo & Domingue, 2018).

Third, using PGIs in social research is not aimed for prediction in the same sense as some health and disease research. Despite the increasing precision of PGIs, for complex behavioural traits such as education, they tend to explain limited proportion of variance in the target traits (compared with PGIs for height and BMI, for instance). Rather, genetically informed social research has been combining PGIs with various designs (family based or unrelated individuals) to help elucidate social mechanisms, such as genetic nurture, assortative mating, and gene-environment correlations which are essentially consequential environmental and socially based mechanisms or contexts. For example, using PGI for education of a large sample of parent-offspring trios in Norway, Nivard et al. (2022) find that indirect genetic effects on children's educational attainment are primarily explained by components derived from multi-generational process of social stratification. Using within-family analysis, J. M. Fletcher et al. (2023) show that the association between EA PGI and education is

relatively weaker for the sibling that has higher PGI (compared with the other ones), which might indicate that compensatory processes are present for children with lower PGI. As one example for gene-environment correlation (rGE), Abdellaoui et al. (2019) identify genetically related geographical mobility that suggest rGE. They find that individuals who are genetically predisposed to higher educational attainment in Great Britain were more likely to migrate out of lower SES areas such as coal mining areas.

4.4   Sensitivity power analysis

Similar as the analysis for MZ twins, for the DZ twin sample (total N = 3168), the effective sample size is chosen as 2030, which is calculated based on an ICC of 0.56 (the ICC for EA PGI in the DZ twin sample, which is very close to the twin-pair ICC for education related polygenic index reported in Selzam et al. (2019)). The halved sample size is 1584. I use an alpha level of 0.05 with two-sided t tests to plot the relationship between varying effect sizes and statistical power.

The results in Figure A.9 show that, even under the halved sample size, there is over 95% power to detect an effect size of 0.1 (beta). Therefore, the DZ twin sample should have fairly decent power to detect a small effect size. In the present DZ twin study, one sd of EA PGI corresponds to an increase of immigration attitudes scale by about 0.13-0.14 sd (equivalent to beta). So far, there is little research examining the effects of EA PGI on political attitudes, but we can compare these effect sizes with Dawes et al. (2021) which examines the effects of EA PGI on political participation. Dawes et al. (2021) shows that one sd of EA PGI increases individual average turnout across national elections by about 0.01-0.02 (with various samples), and increases the self-reported voting scale by about 0.15-0.18 sd.
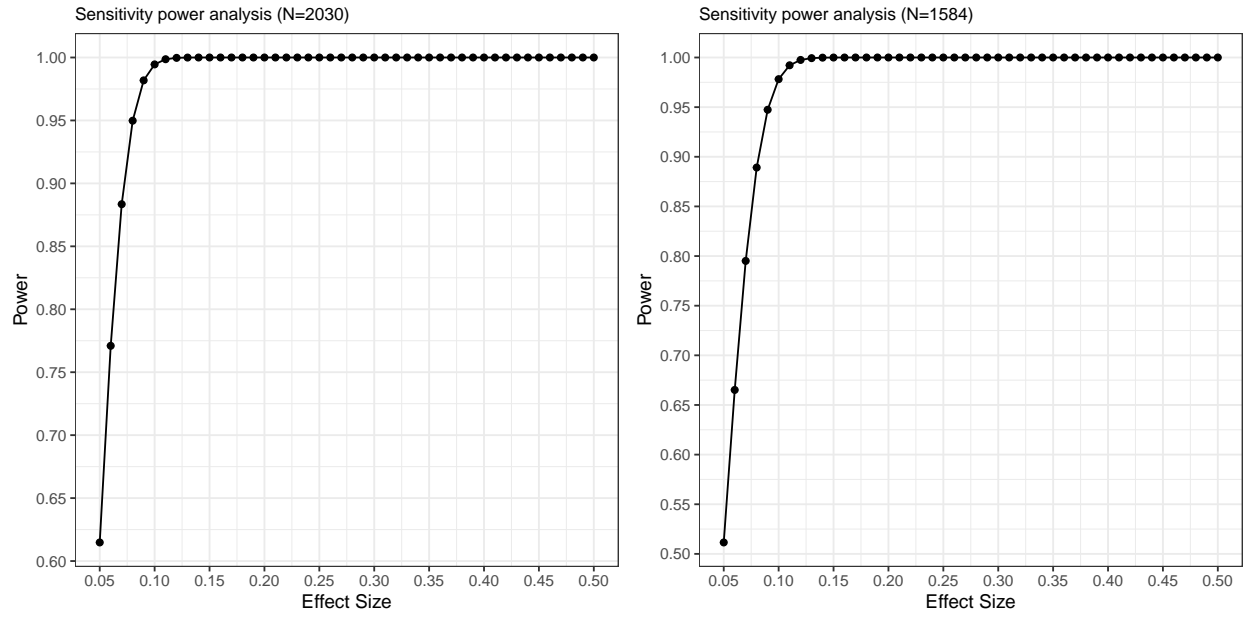
Figure A.9: Sensitivity power analysis for the DZ twin Study

4.5 Full regression results

The effects of EA PGI and education on immigration attitudes (corresponds to Figure 4 in the main text)

Table A.17: The effects of EA PGI and education on immigration attitudes (DZ twins)

VARIABLES	(1) Avg index	(2) Avg index	(3) Avg index	(4) IRT index	(5) IRT index	(6) IRT index	(7) Education
Years of Education	0.087***		0.083***	0.084***		0.079***	
	(0.011)		(0.011)	(0.011)		(0.011)	
EA PGI		0.133***	0.095**		0.139***	0.103**	0.455***
		(0.034)	(0.033)		(0.034)	(0.034)	(0.071)
Constant	-1.021***	0.009	-0.964***	-0.989***	0.005	-0.926***	11.795***
	(0.133)	(0.024)	(0.135)	(0.136)	(0.024)	(0.138)	(0.056)
Observations	3,168	3,168	3,168	3,168	3,168	3,168	3,168
R-squared	0.631	0.620	0.633	0.626	0.616	0.628	0.729
Sex	YES	YES	YES	YES	YES	YES	YES
Twin FE	YES	YES	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at twin-pair level. *** p<0.001, ** p<0.01, * p<0.05.

The effects of EA PGI and education by attitude items

Table A.18: The effects of EA PGI and education by attitude items (DZ twins)

VARIABLES	(1) refugee	(2) refugee	(3) refugee	(4) labor	(5) labor	(6) labor	(7) culture	(8) culture	(9) culture	(10) language	(11) language	(12) language
Years of Education	0.088***		0.083***	0.078***		0.073***	0.044***		0.045***	0.078***		0.070***
	(0.014)		(0.014)	(0.012)		(0.012)	(0.012)		(0.012)	(0.015)		(0.015)
EA PGI		0.134**	0.096*		0.127***	0.094**		0.001	-0.019		0.174***	0.142***
		(0.042)	(0.041)		(0.036)	(0.036)		(0.035)	(0.035)		(0.041)	(0.041)
Constant	1.748***	2.784***	1.806***	1.705***	2.625***	1.762***	1.654***	2.170***	1.642***	1.512***	2.428***	1.597***
	(0.171)	(0.030)	(0.173)	(0.144)	(0.026)	(0.146)	(0.137)	(0.025)	(0.138)	(0.174)	(0.031)	(0.176)
Observations	3,168	3,168	3,168	3,168	3,168	3,168	3,168	3,168	3,168	3,168	3,168	3,168
R-squared	0.608	0.601	0.610	0.585	0.577	0.587	0.577	0.573	0.577	0.575	0.572	0.578
Sex	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Twin FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at twin-pair level. Dependent variables are on the original 1-5 scale. *** p<0.001, ** p<0.01, * p<0.05.

Between family analysis of EA PGI, education, and immigration attitudes

Table A.19: The effects of EA PGI and education on immigration attitudes (Between family)

VARIABLES	(1) Avg index	(2) Avg index	(3) Avg index	(4) IRT index	(5) IRT index	(6) IRT index	(7) Education
Years of Education	0.109***		0.098***	0.115***		0.103***	
	(0.007)		(0.007)	(0.007)		(0.007)	
EA PGI		0.170***	0.095***		0.188***	0.109***	0.766***
		(0.018)	(0.019)		(0.018)	(0.018)	(0.048)
Constant	-1.321***	-0.064	-1.193***	-1.371***	-0.038	-1.224***	11.488***
	(0.128)	(0.099)	(0.130)	(0.130)	(0.101)	(0.132)	(0.275)
Observations	3,168	3,168	3,168	3,168	3,168	3,168	3,168
R-squared	0.109	0.057	0.117	0.118	0.063	0.129	0.119
Controls	YES	YES	YES	YES	YES	YES	YES
Twin FE	NO	NO	NO	NO	NO	NO	NO

Notes: All models include controls for sex, birth-year dummies, interaction between sex and the birth-year dummies, the first twenty principal components of the genetic relationship matrix, and batch fixed effects. Standard errors are clustered at twin-pair level. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

4.6 A simple test on mediators

To explore other potential mediators between EA PGI and immigration attitudes, I conducted some simple test with the following criteria: EA PGI should have significant effect on the mediator, and the mediator should plausibly affect immigration attitudes beyond education. The results reported here should be interpreted as only suggestive.

I focus on a psychological trait - locus of control (LOC) from available data. The LOC classifies individuals along a single dimension capturing the degree to which they feel like they control the outcome of events. Those with an internal locus of control believe they have command over their own fate, attributing achieved outcomes to their own hard work and abilities. Conversely, individuals with an external locus of control perceive outcomes as beyond their control. Individual LOC has been shown to predict immigration sentiments - those who feel in control are less hostile to immigrants (Harell et al., 2017). There has also been extensive literature showing the connection between LOC and academic achievements (e.g., Findley & Cooper (1983)). The LOC used is based on a 12-item forced-choice questionnaire (Rotter, 1966), and the responses to each choice set are summed together. LOC is reversely coded so that higher scores are associated with higher internal locus of control. The variable is standardised.

The first column in the table below shows that EA PGI has significant effects on LOC. Column 2 confirms that higher personal control has positive effects on pro-immigration attitudes, while column 5 shows that LOC slightly attenuates the effects of EA PGI by 6%. Including education decrease the estimates for both EA PGI and LOC, but LOC remains a significant predictor for immigration attitudes. Comparing column 4 and 6 where education has been controlled for, the inclusion of LOC attenuates the effects of EA PGI by about 3%. The patterns for the IRT index are largely identical to those for the average index.

Table A.20: The effects of EA PGI and Locus of control on immigration attitudes (DZ twins)

VARIABLES	(1) LOC	(2) Avg index	(3) Avg index	(4) Avg index	(5) Avg index	(6) Avg index	(7) IRT index	(8) IRT index	(9) IRT index	(10) IRT index	(11) IRT index
EA PGI	0.076*		0.136**	0.095*	0.128**	0.092*		0.139***	0.100*	0.132**	0.097*
	(0.038)		(0.041)	(0.041)	(0.041)	(0.040)		(0.042)	(0.041)	(0.042)	(0.041)
LOC		0.100**			0.094**	0.075*	0.096**			0.090**	0.072*
		(0.031)			(0.030)	(0.030)	(0.030)			(0.030)	(0.030)
Years of Education				0.084***		0.080***			0.082***		0.078***
				(0.014)		(0.014)			(0.014)		(0.014)
Constant	0.090**	-0.011	0.001	-1.002***	-0.008	-0.963***	-0.019	-0.008	-0.979***	-0.016	-0.941***
	(0.028)	(0.027)	(0.027)	(0.166)	(0.027)	(0.164)	(0.028)	(0.028)	(0.168)	(0.027)	(0.167)
Observations	2,296	2,296	2,296	2,296	2,296	2,296	2,296	2,296	2,296	2,296	2,296
R-squared	0.575	0.611	0.611	0.624	0.615	0.627	0.609	0.610	0.622	0.613	0.624
Sex	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Twin FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Standard errors are clustered at twin-pair level. LOC is a standardised variable based on a 12-item battery. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

Locus of Control items

For each pair of statement choose the one that best describes how you feel.
1. Many of the unhappy things in people's lives are partly due to bad luck.
2. People's misfortunes result from the mistakes they make.

1. An important reason why there are wars is that people are not interested enough in politics.
2. There will always be war no matter what many efforts people make to prevent it.

1. In the long run, people get the respect they deserve in this world.
2. Unfortunately, an individual's worth often passes unrecognized no matter how hard he tries.

1. The idea that teachers are unfair to students is nonsense.
2. Most students don't realize the extent to which their grades are affected by accidental happenings.

1. Without the right luck, one cannot be an effective leader.
2. Capable people who fail to became leaders have not taken advantage of their opportunities.

1. No matter how hard you try, some people just don't like you.
2. People who can't get others to like them don't understand how to get along with others.

1. What happens, happens.
2. Trusting fate has never worked as well for me as deciding on a particular plan of action.

1. Becoming a success is a matter of hard work; luck has little or nothing to do with.
2. Getting a good job depends mainly on being in the right place at the right time.

1. An average citizen can influence government decisions.
2. The world is ruled by a few people in power and there is not much that the little man can do about.

1. When I make plans, I am almost certain that I can make them work.
2. It is not always wise to plan too far ahead because many things turn out to be a matter of luck anyway.

1. In my case, getting what I want has little or nothing to do with luck.
2. Many times we might just as well decide what to do by flipping a coin.

1. What happens to me is my own doing.
2. Sometimes I feel that I don't have enough control over the direction my life is taking.

5 References

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K. E., Nivard, M. G., Veul, L., Holtz, Y., Zietsch, B. P., Frayling, T. M., Wray, N. R., Yang, J., Verweij, K. J. H., & Visscher, P. M. (2019). Genetic correlates of social stratification in great britain. *Nature Human Behaviour*, *3*(12), 1332–1342. https://doi.org/10.1038/s41562-019-0757-5

Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, *110*(3), 512–529. https://doi.org/10.1017/S0003055416000216

Becker, J., Burik, C. A. P., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., Belsky, D. W., Karlsson Linnér, R., Ahlskog, R., Kleinman, A., Hinds, D. A., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., . . . Okbay, A. (2021). Resource profile and user guide of the polygenic index repository. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-021-01119-3

Branigan, A. R., McCallum, K. J., & Freese, J. (2013). Variation in the heritability of educational attainment: An international meta-analysis. *Social Forces*, *92*(1), 109–140. https://www.jstor.org/stable/43287519

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H. L., McClay, J., Mill, J., Martin, J., Braithwaite, A., & Poulton, R. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*, *301*(5631), 386–389. https://doi.org/10.1126/science.1083968

Cavaille, C., & Marshall, J. (2019). Education and anti-immigration attitudes: Evidence from compulsory schooling reforms across western europe. *American Political Science Review*, *113*(1), 254–263. https://doi.org/10.1017/S0003055418000588

Cawley, J., Han, E., Kim, J., & Norton, E. (2020). *Sibling correlation in educational attainment: A test of genetic nurture.* https://doi.org/10.3386/w27336

Correia, S. (2017). Reghdfe: Stata module for linear and instrumental-variable/GMM regression absorbing multiple levels of fixed effects. *Statistical Software Components S457874, Boston College Department of Economics.*

Correia, S. (2023). *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.* https://EconPapers.repec.org/RePEc:boc:bocode:s457874

Dawes, C. T., Okbay, A., Oskarsson, S., & Rustichini, A. (2021). A polygenic score for educational attainment partially predicts voter turnout. *Proceedings of the National Academy of Sciences*, *118*(50), e2022715118. https://doi.org/10.1073/pnas.2022715118

Dinesen, P. T., Dawes, C. T., Johannesson, M., Klemmensen, R., Magnusson, P., Nørgaard, A. S., Petersen, I., & Oskarsson, S. (2016). Estimating the impact of education on political participation: Evidence from monozygotic twins in the united states, denmark and sweden. *Political Behavior*, *38*(3), 579–601. https://doi.org/10.1007/s11109-015-9328-2

Findley, M. J., & Cooper, H. M. (1983). Locus of control and academic achievement: A literature review. *Journal of Personality and Social Psychology*, *44*(2), 419–427. https://doi.org/10.1037/0022-3514.44.2.419

Finseraas, H., Skorge, Ø. S., & Strøm, M. (2018). Does education affect immigration attitudes? Evidence from an education reform. *Electoral Studies*, *55*, 131–135. https://doi.org/10.1016/j.electstud.2018.06.009

Fletcher, J. M., Wu, Y., Zhao, Z., & Lu, Q. (2023). The production of within-family inequality: Insights and implications of integrating genetic data. *PNAS Nexus*, *2*, 1–7.

Fletcher, J., Wu, Y., Li, T., & Lu, Q. (2021). *Interpreting polygenic score effects in sibling analysis.*

Harell, A., Soroka, S., & Iyengar, S. (2017). Locus of control and anti-immigrant sentiment in canada, the united states, and the united kingdom. *Political Psychology*, *38*(2), 245–260. https://doi.org/10.1111/pops.12338

Herreros, F., & Criado, H. (2009). Social trust, social capital and perceptions of immigration. *Political Studies*, *57*(2), 337–355. https://doi.org/10.1111/j.1467-9248.2008.00738.x

Hivert, V., Sidorenko, J., Rohart, F., Goddard, M. E., Yang, J., Wray, N. R., Yengo, L., & Visscher, P. M. (2021). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *The American Journal of Human Genetics*, *108*(5), 786–798. https://doi.org/10.1016/j.ajhg.2021.02.014

Hjalmarsson, R., Holmlund, H., & Lindquist, M. J. (2015). The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *The Economic Journal*, *125*(587), 1290–1326. https://doi.org/10.1111/ecoj.12204

Holmlund, H. (2007). A researcher's guide to the swedish compulsory school reform. *CEE Discussion Papers.* http://ideas.repec.org/p/cep/ceedps/0087.html

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., Chittoor, G., Ahlskog, R., Lind, P. A., Palviainen, T., Zee, M. D. van der, Cheesman, R., Mangino, M., Wang, Y., Li, S., Klaric, L., Ratliff, S. M., Bielak, L. F., Nygaard, M., . . . Davies, N. M. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592. https://doi.org/10.1038/s41588-022-01062-7

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. https://doi.org/10.1111/ajps.12261

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. https://doi.org/10.1016/j.econedurev.2019.02.002

Magnusson, P. K. E., Almqvist, C., Rahman, I., Ganna, A., Viktorin, A., Walum, H., Halldner, L., Lundström, S., Ullén, F., Långström, N., Larsson, H., Nyman, A., Gumpert, C. H., Råstam, M., Anckarsäter, H., Cnattingius, S., Johannesson, M., Ingelsson, E., Klareskog, L., . . . Lichtenstein, P. (2013). The swedish twin registry: Establishment of a biobank and other recent developments. *Twin Research and Human Genetics*, *16*(1), 317–329. https://doi.org/10.1017/thg.2012.104

Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *The American Economic Review*, *95*(1), 414–424. https://www.jstor.org/stable/4132688

Meghir, C., Palme, M., & Simeonova, E. (2018). Education and mortality: Evidence from a social experiment. *American Economic Journal: Applied Economics*, *10*(2), 234–256. https://www.jstor.org/stable/26528388

Muslimova, D., Dias Pereira, R., Von Hinke, S., Van Kippersluis, H., Rietveld, C. A., & Meddens, S. F. W. (2023). Rank concordance of polygenic indices. *Nature Human Behaviour*, *7*(5), 802–811. https://doi.org/10.1038/s41562-023-01544-6

Nivard, M. G., Belsky, D., Harden, K. P., Baier, T., Andreassen, O. A., Ystrom, E., Van Bergen, E., & Lyngstad, T. H. (2022). *Neither nature nor nurture: Using extended pedigree data to understand indirect genetic effects on offspring educational outcomes* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/bhpm5

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., . . . Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide

association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449. https://doi.org/10.1038/s41588-022-01016-z

Österman, M. (2021). Can we trust education for fostering trust? Quasi-experimental evidence on the effect of education and tracking on social trust. *Social Indicators Research*, *154*(1), 211–233. https://doi.org/10.1007/s11205-020-02529-y

Pekkarinen, T., Uusitalo, R., & Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the finnish comprehensive school reform. *Journal of Public Economics*, *93*(7), 965–973. https://doi.org/10.1016/j.jpubeco.2009.04.006

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., Westra, H.-J., Shakhbazov, K., Abdellaoui, A., Agrawal, A., Albrecht, E., Alizadeh, B. Z., Amin, N., Barnard, J., Baumeister, S. E., Benke, K. S., Bielak, L. F., Boatman, J. A., Boyle, P. A., . . . Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *340*(6139), 1467–1471. https://doi.org/10.1126/science.1235488

Rotter, J. B. (1966). *Psychological monographs: General and applied GENERALIZED EXPECTANCIES FOR INTERNAL VERSUS EXTERNAL CONTROL OF REINFORCEMENT* (1).

Selzam, S., Ritchie, S. J., Pingault, J. B., Reynolds, C. A., O'Reilly, P. F., & Plomin, R. (2019). Comparing within- and between-family polygenic score prediction. *American Journal of Human Genetics*, *105*(2), 351–363. https://doi.org/10.1016/j.ajhg.2019.06.006

Sidanius, J., Laar, C. van, Levin, S., & Sinclair, S. (2003). Social hierarchy maintenance and assortment into social roles: A social dominance perspective. *Group Processes & Intergroup Relations*, *6*(4), 333–352. https://doi.org/10.1177/13684302030064002

Stubager, R. (2008). Education effects on authoritarian-libertarian values: A question of socialization1. *British Journal of Sociology*, *59*(2), 327–350. https://doi.org/10.1111/j.1468-4446.2008.00196.x

Surridge, P. (2016). Education and liberalism: Pursuing the link. *Oxford Review of Education*, *42*(2), 146–164. https://doi.org/10.1080/03054985.2016.1151408

Tavakol, M., & Dennick, R. (2011). Making sense of cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Trejo, S., & Domingue, B. W. (2018). Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses. *Biodemography and Social Biology*, *64*(3), 187–215. https://doi.org/10.1080/19485565.2019.1681257

Van de Werfhorst, H. G., & De Graaf, N. D. (2004). The sources of political orientations in post-industrial society: Social class and education revisited. *British Journal of Sociology*, *55*(2), 211–235. https://doi.org/10.1111/j.1468-4446.2004.00016.x

Wu, Y., Zhong, X., Lin, Y., Zhao, Z., Chen, J., Zheng, B., Li, J. J., Fletcher, J. M., & Lu, Q. (2021). Estimating genetic nurture with summary statistics of multigenerational genome-wide association studies. *Proceedings of the National Academy of Sciences*, *118*(25), e2023184118. https://doi.org/10.1073/pnas.2023184118

Young, A. I. (2019). Solving the missing heritability problem. *PLOS Genetics*, *15*(6), e1008222. https://doi.org/10.1371/journal.pgen.1008222

Young, A. I., Benonisdottir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype-phenotype associations in humans. *Science*, *365*(6460), 1396–1400. https://doi.org/10.1126/science.aax3710

Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., Masson, G., Thorsteinsdottir, U., Stefansson, K., & Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*(9),

1304–1310. https://doi.org/10.1038/s41588-018-0178-9