Can we trust education for fostering trust?
Quasi-experimental evidence on the effect of education and tracking on social trust

Supplementary material

A.1 Data

Table A.1: Descriptive statistics

	count	mean	sd	min	max
Birth year	68796	1959.2	11.64	1926	1993
Age	68796	50.76	11.96	25	80
Years of full-time education	68211	12.66	4.195	0	25
Most people can be trusted or you can't be too careful	68733	5.091	2.391	0	10
Most of the time people helpful or mostly looking out for themselves	68665	4.904	2.279	0	10
Most people try to take advantage of you, or try to be fair	68548	5.727	2.208	0	10
Social trust scale	68796	5.239	1.901	0	10
High parental education	64960	0.342	0.474	0	1
Female	68796	0.528	0.499	0	1
Belongs to ethnic minority	68796	0.0151	0.122	0	1
Foreign-born father, born within Europe	68796	0.0282	0.166	0	1
Foreign-born mother, born within Europe	68796	0.0261	0.160	0	1
Foreign-born father, born outside of Europe	68796	0.0103	0.101	0	1
Foreign-born mother, born outside of Europe	68796	0.00926	0.0958	0	1
General reform indicator	68796	0.491	0.500	0	1
Reform indicator: extension reforms	40632	0.514	0.500	0	1
Reform indicator: all detracking reforms	29931	0.447	0.497	0	1
Reform indicator: pures<detracking reforms	8258	0.428	0.495	0	1
Understand different people	66903	7.317	2.056	0	10
Institutional trust index	59255	4.491	1.993	0	10

Same sample as for models on social trust in Table 3 in the article, weighted using ESS design weights.

Social trust measure

The scale for social trust is constructed using the following three items from the ESS:

- *ppltrst*: 'Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?'
- *pplfair*: 'Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?'
- *pplhelp*: 'Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves?'

All questions were measured on an 11-point scale ranging from 0 ('You can't be too careful', 'Most people would try to take advantage of me', 'People mostly look out for themselves') to 10 ('Most people can be trusted', 'Most people would try to be fair',

'People mostly try to be helpful'). The scale was then calculated as the mean of the three items, including respondents that had answered at least two of the three questions. Cronbach's alpha for the three-item scale equals 0.77.

The distribution of the social trust indicator in the full sample is shown in Figure A.1 and for the subsamples of respondents with low and high parental education in Figure A.2.

Figure A.1: Distribution of social trust indicator

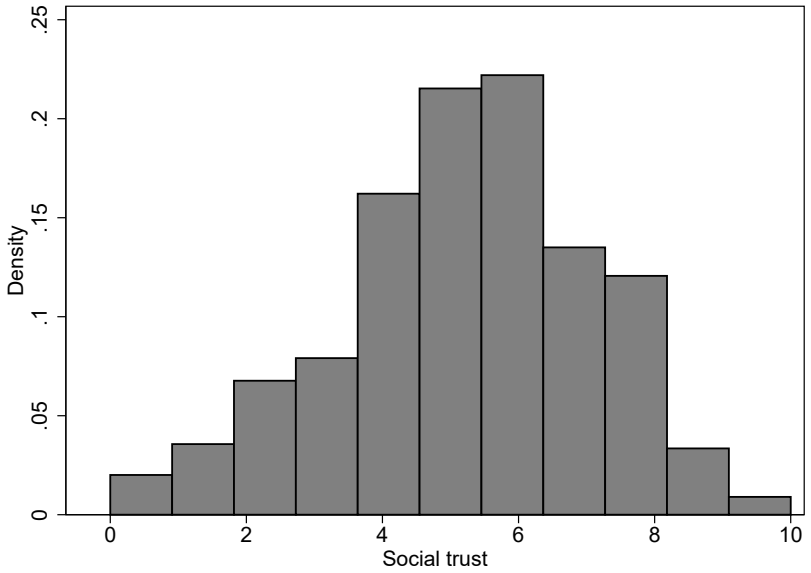
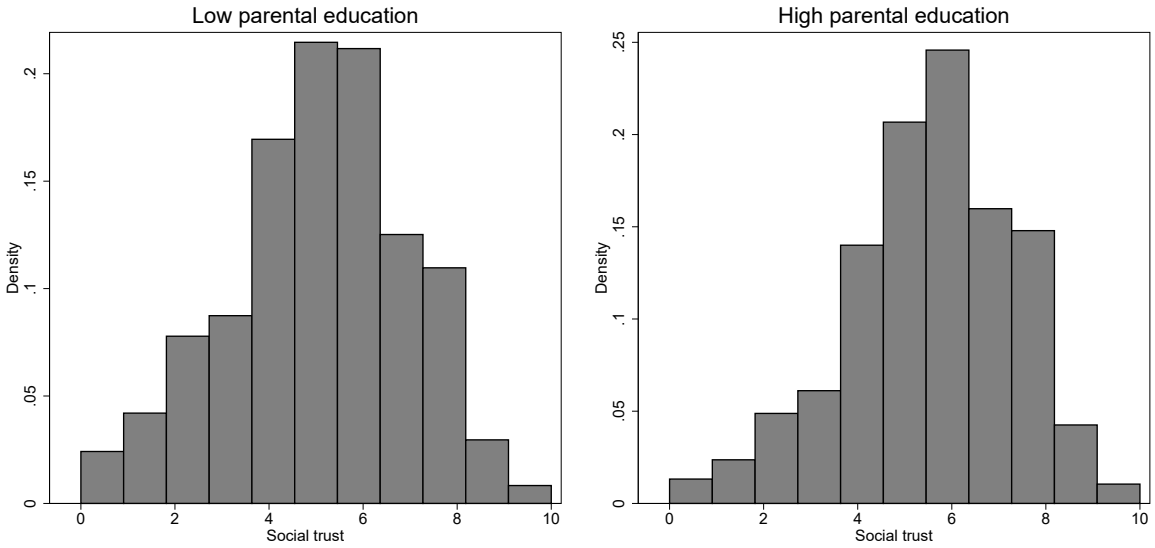


Figure A.2: Distribution of social trust indicator, separated by low and high parental education



Note: The sample is splitted using the same relative measure of parental education as used in the article.

Notes on data

- Years of full-time education. As there are a few instances of unreasonably high values for this variable, it is capped at 25 years.

- High parental education. As described in the article, this variable represents whether the average education of the respondent's parents belongs to the upper half of the distribution of average parental education relative to country and birth year. However, due to that a large majority of the parents of the birth cohorts affected by the reforms have an education equal to the lowest level in the ESS; the share of respondents with highly educated parents is considerably less than 50 percent.

- The variable 'Belongs to ethnic minority' is a dummy variable based on the following item in the ESS: 'Do you belong to a minority ethnic group in [country]?' Where [country] is the country where the ESS survey is conducted.

- The variable 'Understand different people' used in Table 4 is based on the item *ipudrst* in the ESS which is part of several questions where the respondent is asked to say whether they find themselves to be like a person with some stated characteristics. It is phrased in the following way: 'Now I will briefly describe some people. Please listen to each description and tell me how much each person is or is not like you. Use this card for your answer. It is important to her/him to listen to people who are different from her/him. Even when she/he disagrees with them, she/he still wants to understand them.' The item uses a 6 point response scale but this is rescaled to run between 0 and 10 to be comparable to the social trust index. The responses are also reverse coded such that higher values mean that the respondent agrees with being like the described person.

- The institutional trust index is an additive index consisting of five ESS items measuring trust in the country's parliament, legal system, police, politicians and political parties (*trstprl, trstlgl, trstplc, trstplt* and *trstprt*). It runs from 0 to 10. This is the same measure as used by Hakhverdian and Mayne (2012). The index is calculated for respondents that have answered at least two of the five questions.

A.2 Educational reforms and exogenous variation

As was briefly discussed in the article, the exogenous character of educational reforms could be questioned by pointing to the fact that reforms are not decided on and implemented randomly. The implementation of reforms may depend on a number of factors such as national norms and government partisanship that, in turn, may be related to social trust levels. In addition, social trust levels in a country could also in itself affect the likelihood of educational reform. I will here expand on the argument in the article why this issue should be a minor problem for the design and the opportunities for making causal inferences. The key is understanding how the reforms are exploited in the article, as well as how they were implemented. The essential point is that although the educational reforms as such are not exogenous, they induced a variation in education that is effectively exogenous for those affected by the reforms.

The empirical approach in the article builds on the fact that educational reforms in general come into effect starting with a certain birth year cohort; thus, whereas those born one year follow the old educational pathway, those born the years thereafter follow the reformed system. By then comparing adjacent birth cohorts that should be similar in all other respects than reform exposure, it is possible to evaluate the impact of a reform on outcomes. This is effectively a natural experiment, comparing outcomes before and after the reform treatment. The reform-fixed effects are important since they absorb all of the between-reform variation and I am thus only using the difference between unaffected and affected birth cohorts within each particular reform-window. This approach means that all general between-reform differences are factored out. Differences in pre-reform levels of social trust should for instance not affect the estimates as these are absorbed by the reform-fixed effects.

Using the within-reform variation implies that only confounders that impact the reform-affected and reform-unaffected birth cohorts differently would risk to bias the estimates. Put differently, for the approach to work for the identification of a causal effect, it is critical that there are no other systematic differences between the cohorts apart from differences in reform exposure. If reform implementation coincides with a sudden political change, such as a new government coming into power, it could possibly bring this assumption into question, as this change may affect outcomes. However, I assess this risk as small, considering the empirical setup. First, for most reforms, there is a considerable time period between the reform decision and reform implementation. This period could be anything from a year up to ten years. Consequently, the timing of the reform decision usually differs substantially from when school children start to be affected by the reform. A sudden political change connected to the reform decision would seldom coincide with the actual reform implementation. Second, in most cases, I would deem sudden changes of attitudes such as social trust directly connected to government turnover as unlikely. Such attitudinal changes would rather be part of continuously ongoing trends. As the model specification in the article controls for time trends over cohorts, such long-term effects should largely be absorbed by the controls. Changes in attitudes over time would effectively only be an issue if there is a sudden change in direct connection to reform implementation, which again differs from the actual political decision on the reform. Third, most of the reforms would affect children at rather young ages, often between 10 to 15 years old. To the degree that these reforms would coincide with government turnovers or other sudden political changes that could affect attitudes, you would expect these changes to mostly affect older children or perhaps mainly the parents. Naturally, the

parents could in turn have an impact on their children. However, it appears unlikely that the degree to which parents would be affected would be strongly dependent on whether their own children are part of the cohorts that will be exposed to the reforms. In other words, you would expect both cohorts affected and unaffected by a reform to be exposed to such possible changes in attitudes, which would largely mitigate the issue from a causal identification point of view.

Could there be a risk of reverse causation, meaning that the results on social trust actually depend on differences in social trust affecting the implementation of the reforms? I assess reverse causality in the strict sense as almost impossible with the present design. This would be reverse causality in the sense that the effect of the detracking reforms is a result of differences in social trust prior to the reforms, differences that also, in turn, explain who becomes affected by the reforms. However, such a causal link is hardly imaginable. First, as stated above, any before-reform differences in social trust are absorbed by the reform-fixed effects. The difference in effect between extension and detracking reforms could thus not be explained by prior differences in social trust levels. Reversed causality would instead with the present research design imply something in line with that parents with higher levels of social trust would postpone getting children in order to assure that their children would go to the new detracked school, whereas this would not happen for extension reforms. This seems highly unlikely and would also for most of the reforms be impossible as the decisions about implementing the reforms primarily were taken at a time when the to-be-affected children already were born. From a less strict understanding of reversed causality the empirical design could be criticized by arguing that educational reforms are more likely to be carried through in countries with relatively high levels of social trust. However, such an argument is strictly not an argument about reversed causality, but rather an issue of generalization. In other words, it is a question of external rather than internal validity. That is, if social trust levels affect the likelihood of educational reform, the sample of reforms have an overrepresentation of high trust countries. Such an overrepresentation means that it could be questioned whether the effect of the reforms could be generalized to low-trust countries. While this is a valid objection, it is not a question of whether there actually is a causal effect of detracking reforms on social trust; it is only a question of whether we could expect a similar effect in countries that have not yet carried through such reforms. On this latter issue I would like to underline that the sample of countries that have implemented detracking reforms is relatively broad, including several Southern European countries with lower levels of social trust than typical high-trust countries in Northern Europe. The data include detracking reforms implemented in Belgium, Denmark, Spain, Finland, France, Greece, Italy, The Netherlands, Poland, Portugal and Sweden.

However, whereas I assume that the reforms have generated exogenous variation in education between unaffected and affected cohorts, what countries that carried through what type of reform is not a matter of random assignment. That is, it is possible that the countries that implemented extension reforms differ systematically from countries that carried through detracking reforms and that these differences influence the different effects of the reforms. My comparison of these two types of reforms thus hinges on the assumption that the set of countries that have carried through these different types of reforms are comparable, and thus that they represent counter-factual cases. It becomes particularly important that there are no systematic differences between these countries in how we would expect education to affect social trust. I argue that this assumption of comparability across types of reforms is reasonable. First, the set of countries that have

carried through extension reforms and detracking reforms, respectively, do not correspond to any typical groups of European countries. Both groups are instead rather diverse and include both Northern, Southern and Eastern European countries. Some countries have actually implemented both types of reforms. Second, an argument in the literature is that the effect of education on social trust is conditioned by the quality of public institutions (Charron and Rothstein 2016) — what is often referred to as *the quality of government* (QoG). Whether my results could be explained this way is explored in section A.8. In short, there is no support for that systematic differences in QoG would explain the different effects of detracking and extensions reforms.

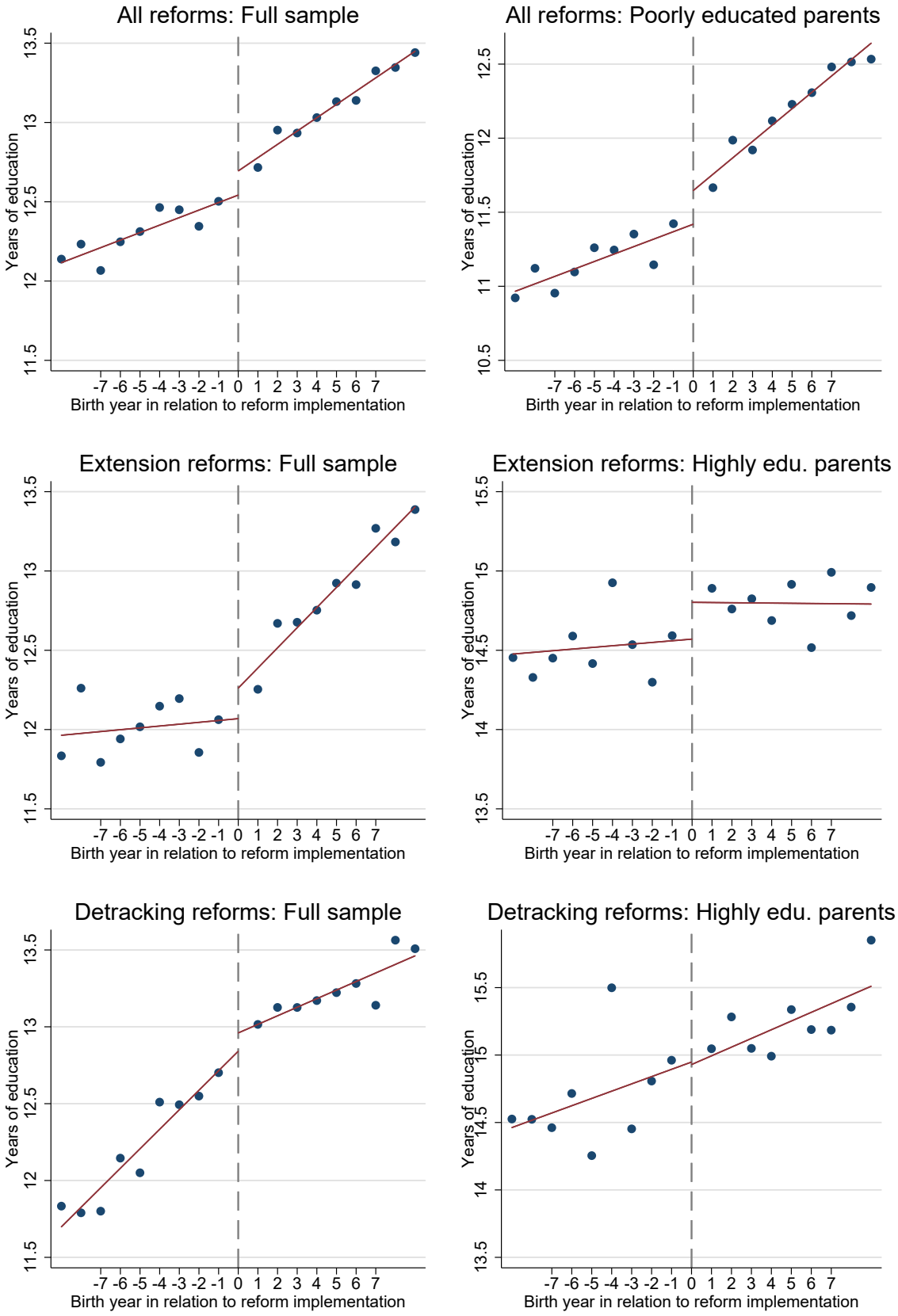
A.3 Reform effect on educational attainment

Figure A.3 offers an overview of the difference in years of education for the treated and non-treated cohorts, with separate graphs for the different reforms and for parental education. The graphs should be interpreted with caution, as they do not include any fixed effects or other controls. Nevertheless, they do reveal some general patterns in the data. A rather clear discontinuity between the treated and the non-treated cohorts is discernible for both the full sample including all reforms and for the extension reforms. For the detracking reforms it is harder to see an effect, especially for respondents with highly educated parents. The patterns also differ between respondents with highly and poorly educated parents. In the case of the extension reforms, their implementation seems to be related to the start of a clear trend of increasing education for those with poorly educated parents. This is likely an effect resulting from some lag in the implementation of some of the reforms that mostly affected this group. In the case of the detracking reforms, almost the opposite seems to be taking place for those with poorly educated parents. A jump upward in educational attainment after the reforms seems to be followed by a shift to a weaker trend. This could possibly be explained by the fact that I for some of these reforms had to approximate the timing of regionally implemented reforms such that a part of the reform effect might kick in before year null (i.e. Sweden and Finland).

Table A.2 shows the effect of the reforms on educational attainment in years of education. Only the dependent variable differs compared to the models in Table 3 in the article. In an Instrumental Variable (IV) framework these models would be the equivalent of the first stage.

The average reform effect for all types of reforms equals 0.13 additional years of education. This effect is twice as large for respondents with poorly educated parents. As could be expected, the effect is also stronger for extension reforms than for detracking reforms. The effect of 0.31 additional years of education for the extension reforms for the group with poorly educated parents is comparable to the effect of 0.30–0.40 years found in papers using a similar set of reforms (Brunello et al. 2009; d'Hombres and Nunziata 2016). The effect of the detracking reforms is smaller and falls just below statistical significance for respondents with poorly educated parents, whereas the effect of pure detracking reforms is close to zero and far from significant.

Figure A.3: Average years of education before and after reform



Average years of education for the seven birth cohorts before and after the first affected cohort (0). The sample is the same as for the regression models. Two separate bivariate regression lines are fitted on the underlying data.

Table A.2: Reform effect on educational attainment (corresponding to the first stage in an IV-framework). Same type of models as in Table 3 in the article.

	All reforms			Extension reforms		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.129**	0.270***	0.133*	0.167**	0.310***	0.110
	(0.059)	(0.067)	(0.073)	(0.066)	(0.079)	(0.079)
High parental edu		2.698***	-130.291		2.712***	-48.113
		(0.079)	(183.076)		(0.105)	(237.056)
Ref × High par edu		-0.189*	0.146		-0.267**	0.304*
		(0.105)	(0.131)		(0.133)	(0.161)
Adj. R^2	0.23	0.31	0.32	0.29	0.37	0.38
Countries	16	16	16	11	11	11
Observations	68266	64474	64474	40334	38245	38245
Flexible interaction	–	No	Yes	–	No	Yes

	Detracking reforms			Pure detracking reforms		
	(7)	(8)	(9)	(10)	(11)	(12)
Reform	0.038	0.183	0.102	-0.072	0.051	0.034
	(0.107)	(0.114)	(0.141)	(0.149)	(0.167)	(0.204)
High parental edu		2.765***	11.495		2.117***	-124.038
		(0.118)	(287.725)		(0.138)	(525.479)
Ref × High par edu		-0.106	-0.013		0.059	0.162
		(0.164)	(0.214)		(0.195)	(0.313)
Adj. R^2	0.11	0.21	0.22	0.082	0.17	0.17
Countries	10	10	10	3	3	3
Observations	29684	27926	27926	8222	7781	7781
Flexible interaction	–	No	Yes	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

A.4 Full tables for the main results

The tables below include the same models as Table 3 and Table 4 in the article but report all of the coefficients for the control variables, except for the age and birth year controls. Since the latter variables are interacted with all of the reform fixed effects, they are not interpretable as single coefficients.

Table A.3: Reform effect on social trust index, including controls variables: All reforms and extension reforms

	All reforms			Extension reforms		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.063**	0.083**	0.077*	0.053	0.067	0.043
	(0.030)	(0.034)	(0.041)	(0.042)	(0.047)	(0.055)
High parental edu		0.340***	112.096		0.336***	111.664
		(0.022)	(91.212)		(0.031)	(123.696)
Ref × High par edu		-0.043	-0.030		-0.045	0.039
		(0.033)	(0.059)		(0.045)	(0.081)
Female	0.061***	0.063***	0.039**	0.050**	0.058***	0.029
	(0.016)	(0.016)	(0.019)	(0.021)	(0.022)	(0.027)
Ethnic minority	-0.261***	-0.226***	-0.200**	-0.283***	-0.231**	-0.191*
	(0.067)	(0.069)	(0.080)	(0.087)	(0.090)	(0.105)
Foreign-born father, Europe	-0.090*	-0.084*	-0.070	-0.073	-0.069	-0.090
	(0.048)	(0.048)	(0.060)	(0.058)	(0.057)	(0.070)
Foreign-born mother, Europe	-0.092*	-0.103**	-0.174***	-0.099	-0.104	-0.174**
	(0.047)	(0.049)	(0.065)	(0.061)	(0.064)	(0.083)
Foreign-born father, outside Europe	-0.053	-0.040	0.045	0.010	0.020	0.165
	(0.078)	(0.079)	(0.124)	(0.133)	(0.136)	(0.200)
Foreign-born mother, outside Europe	-0.087	-0.103	-0.208*	-0.023	-0.035	-0.127
	(0.081)	(0.080)	(0.123)	(0.118)	(0.114)	(0.179)
ESS round 2	0.052	0.025	0.042	0.070	0.036	0.082
	(0.045)	(0.047)	(0.060)	(0.059)	(0.060)	(0.075)
ESS round 3	0.148*	0.107	0.072	0.171	0.129	0.118
	(0.088)	(0.088)	(0.109)	(0.112)	(0.113)	(0.141)
ESS round 4	0.253**	0.216*	0.129	0.297*	0.264*	0.208
	(0.125)	(0.126)	(0.155)	(0.158)	(0.160)	(0.200)
ESS round 5	0.364**	0.322*	0.200	0.407*	0.365*	0.275
	(0.169)	(0.170)	(0.210)	(0.217)	(0.217)	(0.270)
ESS round 6	0.403*	0.342	0.144	0.461*	0.403	0.220
	(0.209)	(0.210)	(0.257)	(0.268)	(0.268)	(0.334)
ESS round 7	0.458*	0.390	0.131	0.529*	0.465	0.259
	(0.246)	(0.248)	(0.307)	(0.314)	(0.316)	(0.399)
ESS round 8	0.671**	0.584**	0.303	0.813**	0.726*	0.481
	(0.288)	(0.289)	(0.354)	(0.372)	(0.372)	(0.458)
ESS round 9	0.835**	0.735**	0.408	0.946**	0.862**	0.592
	(0.327)	(0.327)	(0.407)	(0.420)	(0.418)	(0.527)
Adj. R^2	0.20	0.21	0.21	0.17	0.18	0.18
Countries	16	16	16	11	11	11
Observations	68796	64960	64960	40632	38520	38520
Flexible interaction	–	No	Yes	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.
The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
All models include reform FEs, a general quadratic birth year trend, and reform-specific trends for birth year (linear) and age (quadratic).
The dummies on whether the respondent's parents are foreign-born should be interpreted in relation to the reference category: parents being born in current country of residence. The ESS round dummies should be interpreted in relation to the first round.

Table A.4: Reform effect on social trust index, including controls variables: detracking reforms

	Detracking reforms			Pure detracking reforms		
	(7)	(8)	(9)	(10)	(11)	(12)
Reform	0.082**	0.113**	0.124**	0.030	0.086	0.130*
	(0.040)	(0.046)	(0.059)	(0.055)	(0.058)	(0.073)
High parental edu		0.346***	160.972		0.397***	261.783
		(0.030)	(104.150)		(0.052)	(179.681)
Ref × High par edu		-0.040	-0.091		-0.020	-0.135
		(0.046)	(0.086)		(0.079)	(0.101)
Female	0.059***	0.056***	0.042	-0.004	0.015	-0.011
	(0.022)	(0.021)	(0.027)	(0.035)	(0.035)	(0.046)
Ethnic minority	-0.243**	-0.226**	-0.237**	-0.394**	-0.404**	-0.396*
	(0.099)	(0.103)	(0.119)	(0.172)	(0.192)	(0.204)
Foreign-born father, Europe	-0.134	-0.124	-0.028	-0.196*	-0.153	-0.155
	(0.082)	(0.090)	(0.115)	(0.105)	(0.110)	(0.147)
Foreign-born mother, Europe	-0.070	-0.092	-0.170*	-0.094	-0.048	-0.064
	(0.073)	(0.076)	(0.101)	(0.094)	(0.091)	(0.127)
Foreign-born father, outside Europe	-0.108	-0.095	-0.059	-0.159	-0.138	-0.189
	(0.084)	(0.084)	(0.132)	(0.104)	(0.101)	(0.151)
Foreign-born mother, outside Europe	-0.072	-0.086	-0.125	-0.063	-0.097	-0.032
	(0.104)	(0.103)	(0.153)	(0.141)	(0.136)	(0.157)
ESS round 2	0.028	0.009	-0.021	0.059	0.010	-0.052
	(0.070)	(0.073)	(0.092)	(0.129)	(0.127)	(0.165)
ESS round 3	0.125	0.083	0.011	0.214	0.140	-0.022
	(0.128)	(0.130)	(0.159)	(0.215)	(0.208)	(0.266)
ESS round 4	0.189	0.139	0.018	0.484	0.402	0.134
	(0.189)	(0.192)	(0.225)	(0.302)	(0.296)	(0.331)
ESS round 5	0.320	0.266	0.113	0.485	0.385	0.122
	(0.250)	(0.256)	(0.306)	(0.378)	(0.377)	(0.425)
ESS round 6	0.330	0.248	0.036	0.557	0.399	-0.022
	(0.306)	(0.311)	(0.369)	(0.481)	(0.475)	(0.528)
ESS round 7	0.381	0.287	-0.036	0.673	0.493	-0.048
	(0.363)	(0.370)	(0.439)	(0.567)	(0.556)	(0.626)
ESS round 8	0.517	0.409	0.095	0.848	0.653	0.097
	(0.422)	(0.428)	(0.511)	(0.677)	(0.666)	(0.750)
ESS round 9	0.715	0.570	0.185	1.087	0.852	0.296
	(0.479)	(0.485)	(0.587)	(0.748)	(0.727)	(0.827)
Adj. R^2	0.25	0.26	0.26	0.064	0.079	0.079
Countries	11	11	11	3	3	3
Observations	29931	28153	28153	8258	7808	7808
Flexible interaction	–	No	Yes	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.
The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
All models include reform FEs, a general quadratic birth year trend, and reform-specific trends for birth year (linear) and age (quadratic).
The dummies on whether the respondent's parents are foreign-born should be interpreted in relation to the reference category: parents being born in current country of residence. The ESS round dummies should be interpreted in relation to the first round.

Table A.5: Testing the mechanism behind the effect of detracking reforms[1]

	Social trust: education controls		Social trust: restricted samples		Understand different people		Inst. trust
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reform	0.071*	0.089**	0.154	0.255**	0.081*	0.116**	-0.016
	(0.039)	(0.043)	(0.101)	(0.123)	(0.047)	(0.053)	(0.045)
High parental edu		0.130***				0.323***	
		(0.029)				(0.034)	
Ref × High par edu		-0.029				-0.059	
		(0.045)				(0.052)	
Female	0.068***	0.063***	-0.018	0.100	0.290***	0.280***	-0.075***
	(0.021)	(0.021)	(0.039)	(0.065)	(0.027)	(0.027)	(0.024)
Ethnic minority	-0.214**	-0.219**	-0.486***		-0.059	-0.042	-0.261**
	(0.090)	(0.092)	(0.179)		(0.111)	(0.114)	(0.102)
Foreign-born mother, Europe	-0.103	-0.106	0.051		0.186**	0.115	-0.022
	(0.074)	(0.076)	(0.136)		(0.089)	(0.090)	(0.087)
Foreign-born father, Europe	-0.109	-0.123	-0.307**		0.085	0.140	-0.086
	(0.081)	(0.089)	(0.141)		(0.097)	(0.098)	(0.090)
Foreign-born father, outside Europe	-0.128	-0.103	-0.134	0.118	0.234*	0.133	-0.053
	(0.084)	(0.085)	(0.133)	(0.086)	(0.127)	(0.128)	(0.091)
Foreign-born mother, outside Europe	-0.070	-0.084	-0.068	0.094	0.002	0.056	-0.132
	(0.100)	(0.101)	(0.159)	(0.108)	(0.126)	(0.124)	(0.125)
ESS round 2	0.012	-0.010	0.043	0.116	0.009	0.024	-0.094
	(0.069)	(0.072)	(0.131)	(0.222)	(0.089)	(0.097)	(0.446)
ESS round 3	0.049	0.004	0.203	0.539	-0.051	-0.038	0.141
	(0.126)	(0.127)	(0.208)	(0.371)	(0.169)	(0.179)	(0.469)
ESS round 4	0.107	0.052	0.270	0.739	-0.095	-0.091	0.217
	(0.186)	(0.188)	(0.290)	(0.529)	(0.229)	(0.243)	(0.492)
ESS round 5	0.222	0.156	0.520	0.997	-0.282	-0.294	0.149
	(0.247)	(0.252)	(0.387)	(0.706)	(0.306)	(0.324)	(0.539)
ESS round 6	0.198	0.103	0.453	1.064	-0.147	-0.184	0.083
	(0.302)	(0.305)	(0.467)	(0.895)	(0.373)	(0.394)	(0.573)
ESS round 7	0.235	0.121	0.572	1.470	-0.278	-0.310	0.165
	(0.360)	(0.363)	(0.549)	(1.066)	(0.443)	(0.464)	(0.607)
ESS round 8	0.345	0.218	0.974	1.503	-0.293	-0.331	0.328
	(0.418)	(0.421)	(0.639)	(1.233)	(0.517)	(0.541)	(0.664)
ESS round 9	0.505	0.348	1.105	2.031	-0.263	-0.361	0.757
	(0.476)	(0.479)	(0.697)	(1.379)	(0.591)	(0.618)	(0.722)
Adj. R^2	0.28	0.28	0.22	0.15	0.036	0.038	0.27
Countries	10	10	7	10	10	10	10
Observations	29663	27926	8337	3459	28982	27321	25759
Education FEs	Yes	Yes	No	No	No	No	No
Sample	Full	Full	≤40 yrs	Minority	Full	Full	Full

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.

[1] Sample and controls mimic Table 3 in the article with the following alterations. Models (1) and (2) add a full set of dummies for years of education. Model (3) restricts the sample to respondents up to 40 years old, whereas Model (4) only includes respondents belonging to an ethnic minority, respondents being members of a discriminated group or respondents who have at least one foreign-born parent. Models (5) and (6) exchange the dependent variable to a measure of whether the respondent agrees that it is important to understand people who are different from oneself. In Model (7) the dependent variable is instead an index for institutional trust. See the Online Appendix for further details on the dependent variables in Models (5), (6) and (7).

A.5 Instrumental variable estimations

In the main models in the article I choose to not use an instrumental variable (IV) approach as such an approach assumes that the full effect of the reforms would be mediated by educational attainment (measured in years of education). This assumption is referred to as *the exclusion restriction* (Angrist and Pischke 2009). Making such an assumption contradicts my argument that the institutional context of education—and tracking in particular— may be as important for the outcomes of education as the length of education. It would therefore be unsuitable to use an IV approach for all of the models that include detracking reforms, where the effect of the reforms is assumed to at least partly be a result of less tracking rather than extended education. Using IV estimation in this case would bias the estimates of the effect of educational attainment on social trust since also the detracking effect would be included in the educational attainment effect.

However, for the extension reforms the exclusion restriction is more likely to hold since no changes in tracking were implemented. To exclude the possibility that the non-significant effect of extension reforms on social trust in the article would depend on the estimation technique (reduced form estimates), I have in Table A.6 below used the IV-approach for the extension reforms. The model specification is otherwise the same as in the main models in the article. The point estimates become larger, as expected since they are upscaled in relation to the first stage effect on educational attainment, which is less than 1 year (see Table A.2). However, similar to the extension models in Table 3 in the article, the coefficients fail to reach statistical significance. The interpretation of the effects of the extension reforms is thus similar to that of the main models. There are indications of a positive effect on social trust from extending compulsory education but the effect is imprecise and we are not able to distinguish it from possible random variation.

Table A.6: The effect of extension reforms on social trust, IV estimation

	Extension reforms		
	(1)	(2)	(3)
Years of education	0.349	0.224	0.364
	(0.252)	(0.141)	(0.505)
High parental edu		-1.190	0.000
		(2.649)	(.)
Yrs of edu × High par edu		0.063	-0.135
		(0.172)	(0.533)
Adj. R^2	.	0.081	.
Countries	11	11	11
Observations	40299	38245	38245
Flexible interaction	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies. Years of education is instrumented by the reform dummy (used in Table 3 in the article). In the parental education interaction models the interaction is also instrumented by adding the interaction between the reform dummy and parental education as an instrument.

A.6 Additional models on mechanisms

A.6.1 The effect of extension reforms on detracking mechanisms

Table A.7: The effect of extension reforms on detracking mechanisms

	Social trust: education controls		Social trust: restricted samples		Understand different people		Inst. trust
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reform	0.046	0.051	0.137	0.029	0.009	0.032	0.031
	(0.040)	(0.045)	(0.096)	(0.125)	(0.042)	(0.047)	(0.040)
High parental edu		0.131***				0.261***	
		(0.031)				(0.038)	
Ref × High par edu		-0.026				-0.029	
		(0.044)				(0.056)	
Adj. R^2	0.19	0.19	0.11	0.12	0.033	0.037	0.16
Countries	10	10	7	10	10	10	10
Observations	40299	38245	7920	4846	39732	37711	35192
Education FEs	Yes	Yes	No	No	No	No	No
Sample	Full	Full	≤40 yrs	Minority	Full	Full	Full

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

Table A.7 replicates Table 4 in the article but explores the effect of extension reforms rather than detracking reforms as in Table 4. This represents a placebo test on mechanisms. If the effect of detracking reforms on these outcomes is related to intergroup contacts in school, we would not expect similar effects of extension reforms where the level of tracking was kept unchanged. The results are in line with these expectations, the reform effects in Models (1) to (6) are all smaller in Table A.7 than in Table 4 and none of the reform effects are significant.

A.6.2 The effect of detracking reforms on institutional trust

Table A.8: The effect of detracking reforms on institutional trust

	Main models			Restricted samples	
	(1)	(2)	(3)	(4)	(5)
Reform	-0.016	0.054	0.054	-0.101	0.133
	(0.045)	(0.049)	(0.049)	(0.088)	(0.145)
High parental edu		0.408***	0.408***		
		(0.032)	(0.032)		
Ref × High par edu		-0.093*	-0.093*		
		(0.049)	(0.049)		
Adj. R^2	0.27	0.28	0.28	0.31	0.19
Countries	10	10	10	10	10
Observations	25759	24518	24518	6886	2985
Flexible interaction	–	No	No	–	–
Sample	Full	Full	Full	≤40 yrs	Minority

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

In Table A.8 the main models on detracking reforms from Table 3 and the restricted sample models from Table 4 are rerun, using an index for institutional trust instead of social trust as the dependent variable. The institutional trust index includes five ESS items measuring trust in the country's parliament, legal system, police, politicians and political parties (see notes on data earlier in the Appendix). Except the change of the dependent variable the models are specified as the corresponding models in the paper.

These models are run as a form of placebo test regarding the mechanism. If the positive effect of detracking reforms on social trust depends on intergroup contacts in school, rather than for instance curriculum effects, I would expect smaller or no effects on institutional trust as it is less likely that these attitudes would be affected by the composition of your peers in school. The effects of the reforms do not become significant in any of the models and the point estimates vary between being positive and negative. The significant interaction effect between reform exposure and high parental education only implies that the interaction coefficient is significant, the conditional effect of the reforms on institutional trust for those with highly educated parents is not significant.

A.7 Placebo test: Monte Carlo simulations

In the placebo test reform implementation is randomized by a random draw from a uniform distribution where the first potentially affected birth cohort varies between 1933 and 1986. The placebo reforms follow the same pattern as the actual reforms regarding the number and types of reforms generated in each country. I.e. the same number of extension, detracking and pure detracking reforms are generated in each country as in the actual data. It is just that the timing of the reform implementation (i.e. what cohorts were affected) is randomized). In countries where reforms were implemented regionally, placebo reforms are also regionally generated. In countries with several reforms, additional reforms are not allowed to 'overwrite' already generated reforms (i.e. if reforms happen to be generated so that they overlap). For each set of generated reforms the models in Tables A.9 and A.10 are estimated. In comparison to the models in the article, the flexible interaction models are excluded. The generation of reforms is repeated 10000 times and the estimates from each simulation are stored. Tables A.9 and A.10 present the averages for coefficients, standard errors and other statistics from the Monte Carlo simulations.

Figure A.4 shows the distribution of the coefficient estimates from the placebo Monte Carlo simulations of the average effect of detracking reforms on social trust (based on Model (7) in Table 3 in the article). The actual reform effect of 0.0820 is somewhat more than two standard deviations from the mean of the distribution (-0.00700, the sd of the distribution estimates equals 0.0442). The reform effect thus falls outside of the 95 per cent confidence interval. This in turn implies that if the sampling distribution of the effect is approximated by the generated placebo distribution, the probability of by random getting such an effect is equal to 0.0440 ($P(Z > 2.014)$, two-tailed hypothesis).

Figure A.5 instead presents the distribution of the coefficient estimates from the placebo Monte Carlo simulations of the effect of detracking reforms among individuals with poorly educated parents on social trust (based on Model (8) in Table 3 in the article). The actual reform effect of 0.113 for this group is more than 2.5 standard deviations (1 sd equals 0.0483) from the mean of the distribution (-0.0131). That implies that if the sampling distribution of the effect is approximated by the generated placebo distribution, the probability of by random getting such an effect is equal to 0.00903 ($P(Z > 2.61)$, two-tailed hypothesis).

Table A.9: Placebo test: Reform effect on years of education. Averages from 10000 Monte Carlo simulations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	-0.02	-0.01	-0.02	0.00	-0.02	0.01	0.00	0.04
	(0.06)	(0.06)	(0.08)	(0.08)	(0.10)	(0.10)	(0.19)	(0.19)
High par edu		2.70***		2.65***		2.85***		2.53***
		(0.06)		(0.08)		(0.10)		(0.19)
Ref × High p e		-0.02		-0.02		-0.08		-0.08
		(0.07)		(0.09)		(0.10)		(0.16)
Adj. R^2	0.30	0.38	0.32	0.39	0.28	0.36	0.12	0.21
Countries	16	16	11	11	10	10	3	3
Observations	58860	55592	34437	32561	28591	27060	10063	9604
Reforms	All	All	Ext	Ext	Detrack	Detrack	Pure de	Pure de

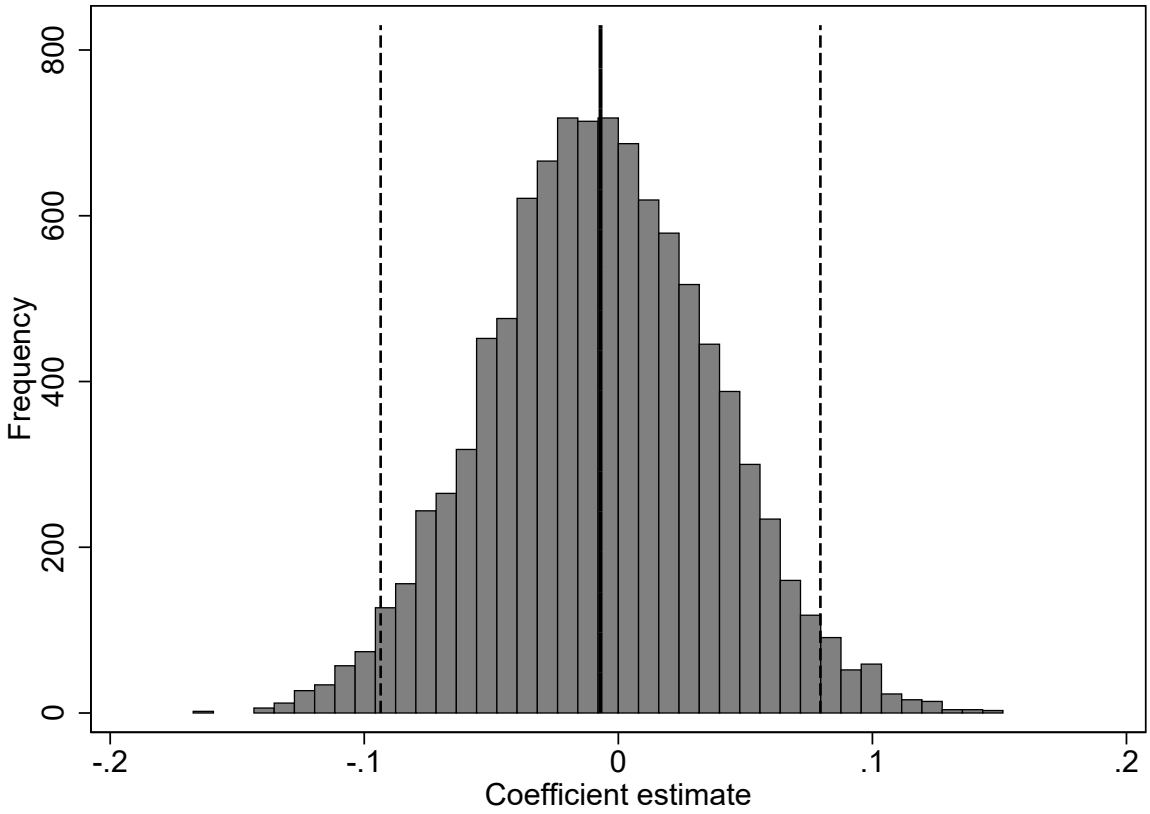
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.
Ext: Extended schooling reforms. Detrack: all detracking reforms. Pure de: detracking only reforms.
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

Table A.10: Placebo test: Reform effect on social trust. Averages from 10000 Monte Carlo simulations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	-0.00	-0.01	0.00	-0.00	-0.01	-0.01	-0.02	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.07)	(0.08)
High par edu		0.32***		0.32***		0.32***		0.33***
		(0.02)		(0.03)		(0.03)		(0.04)
Ref × High p e		0.02		0.02		0.02		0.02
		(0.03)		(0.04)		(0.04)		(0.06)
Adj. R^2	0.20	0.21	0.15	0.16	0.26	0.27	0.10	0.11
Countries	16	16	11	11	11	11	3	3
Observations	64012	60443	37648	35574	30793	29141	11193	10646
Reforms	All	All	Ext	Ext	Detrack	Detrack	Pure de	Pure de

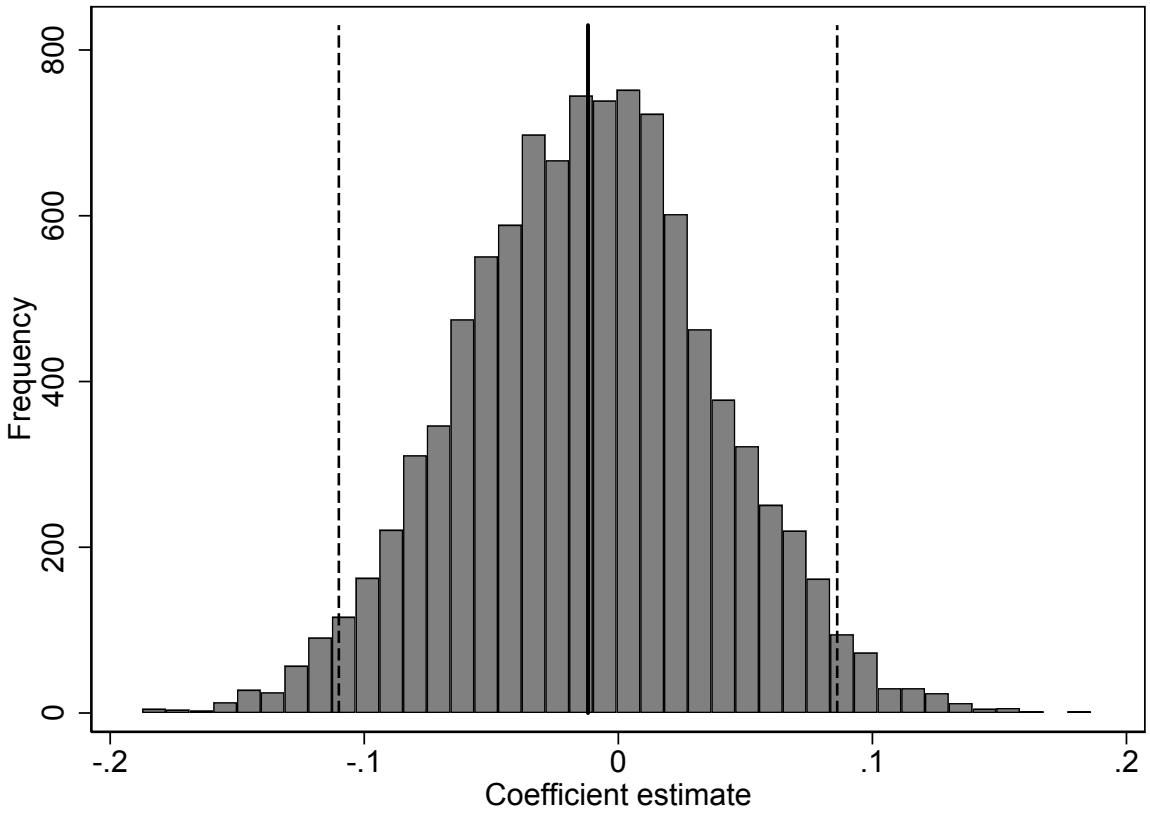
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.
Ext: Extended schooling reforms. Detrack: all detracking reforms. Pure de: detracking only reforms.
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

Figure A.4: Placebo test: effect of fake detracking reforms on social trust. Based on Model (7) in Table 3. Distribution of coefficient estimates from 10000 Monte Carlo simulations.



The solid line denotes the mean (-0.0070) and the dashed lines $+/-1.96$ standard deviations from the mean (1 sd equals 0.044). The dashed lines thus portray the limits for what effects may be expected at random on a 95 per cent confidence level.

Figure A.5: Placebo test: effect of fake detracking reforms on social trust among individuals with poorly educated parents. Based on Model (8) in Table 3. Distribution of coefficient estimates from 10000 Monte Carlo simulations.



The solid line denotes the mean (-0.0131) and the dashed lines $+/-1.96$ standard deviations from the mean (1 sd equals 0.00483). The dashed lines thus portray the limits for what effects may be expected at random on a 95 percent confidence level.

A.8 Influence of quality of government

To test the influence of quality of government (QoG), I construct an QoG indicator closely resembling the measure used by Charron and Rothstein (2016). I use data from the *World Governance Indicators*[4] on 'control of corruption', 'government effectiveness', 'rule of law' and 'voice and accountability' (same indicators as Charron and Rothstein). I take the average of these for each country and use data from the year 2000. The year is chosen on the basis that the data from the first ESS round is from 2002. This general QoG indicator is rescaled to run between 0 and 1.

I start to test the influence of QoG by splitting the sample of countries on the basis of my QoG measure and run models that include all types of reforms. High-QoG countries: Austria, Germany, Denmark, Finland, the UK, Ireland, the Netherlands and Sweden. Low-QoG countries: Belgium, Spain, France, Greece, Hungary, Italy, Poland and Portugal. These models are presented in Table A.11 below. If it would be the case that the difference between extension and detracking reforms rather reflects differences between countries with a low and high level of QoG, we would see a similar pattern as when I separated these different types of reforms in the article. However, the results show that the effect of the reforms on social trust is *stronger* in countries with *a low level of QoG*. Model (1) shows that in low-QoG countries the general positive reform effect is significant on the 95 percent level. When I let the reform effect vary with parental education among countries with low institutional quality in Model (2), the positive effect is significant at the 99 percent level. In high-QoG countries the point estimates are smaller and far from statistically significant.

I continue to assess the role of QoG by interacting the contionous QoG indicator with the reform coefficient to reveal how the effects of the reforms vary with institutional quality. I run separate models for the joint set of all reforms, the extension reforms and

[4]Kaufmann, D., Kraay, A. & Mastruzzi, M., 2011. The worldwide governance indicators: methodology and analytical issues. *Hague J. Rule Law* 3(02): 220–246.

Table A.11: Reform effect on social trust: influence of quality of government. Split sample: low and high QoG.

	All reforms			
	(1)	(2)	(3)	(4)
Reform	0.102**	0.129***	0.022	0.033
	(0.042)	(0.046)	(0.039)	(0.043)
High parental edu		0.385***		0.296***
		(0.031)		(0.028)
Ref × High par edu		-0.058		-0.025
		(0.043)		(0.040)
Adj. R^2	0.076	0.086	0.083	0.091
Countries	8	8	8	8
Observations	36251	34509	32545	30451
Quality of gov	Low	Low	High	High

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

Table A.12: Reform effect on social trust: influence of quality of government. Reform effect interacted with continuous QoG measure.

	All reforms		Extension reforms		Detracking reforms	
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.137*	0.150*	0.162	0.169	0.136	0.134
	(0.073)	(0.079)	(0.102)	(0.109)	(0.109)	(0.113)
Ref × QoG	-0.129	-0.119	-0.203	-0.191	-0.086	-0.039
	(0.102)	(0.110)	(0.162)	(0.167)	(0.135)	(0.140)
High parental edu		0.446***		0.379***		0.458***
		(0.047)		(0.056)		(0.065)
Ref × High par edu		-0.103		-0.042		-0.062
		(0.076)		(0.045)		(0.115)
QoG × High par edu		-0.181**		-0.083		-0.180**
		(0.070)		(0.087)		(0.090)
Ref × QoG × High p edu		0.101				0.032
		(0.110)				(0.152)
Adj. R^2	0.20	0.21	0.17	0.18	0.25	0.26
Countries	16	16	11	11	10	10
Observations	68796	64960	40632	38520	29931	28153

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses.
QoG: Aggregate Quality of Government measure. Rescaled to 0–1.
Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

detracking reforms. The small number of pure detracking reforms does not allow for testing these interactions. I runt two types of models for each set of reforms. First a model where only the reform effect is interacted with the QoG indicator. The second type of model explores whether a possible conditioning effect of QoG could vary with parental education by adding such interactions. However, estimating a number of interaction terms with the limited country-level variation in QoG calls for some caution in the interpretation. Note that there are no coefficients estimated for the QoG constitutive element of the interactions since this variation is absorbed by the reform-fixed effects.

The results are presented in Table A.12. A study of Models (1), (3) and (5) indicates that a high level of QoG is associated with smaller reform effects on social trust. The negative interaction is somewhat larger for extension reforms than for detracking reforms, even though none of the interactions are near significance. Models (2), (4) and (6) that allow the effects of QoG and the reforms to vary with parental education are harder to interpret considering the number of interactions. However, we may focus on the first two coefficients for the reform effect and its interaction with QoG as these portray the effects for the respondents with poorly educated parents. These coefficients demonstrate a similar pattern as the earlier models, a higher level of QoG is associated with a smaller reform effect for both extension and detracking reforms.

The regressions in Table A.12 corroborate the results in Table A.11. If anything, a *higher* level of institutional quality is associated with a *smaller* reform effect on social trust. This holds for both extension and detracking reforms. There is no support for that the stronger positive effect of detracking reforms on social trust is a result of systematic differences in institutional quality between countries that have carried through extension

and detracking reforms. So how come my results are so different from Charron and Rothstein (2016)? Naturally, this is difficult to know with certainty but the main difference is that their data on education is observational whereas mine build on the exogenous variation induced by education reforms. In other words, their educational measure may rather act as a proxy for other unobserved individual-level factors. A number of previous findings on the effects of education have had to be reconsidered recently as more credible research designs have been unable to reproduce these results (e.g., Kam and Palmer 2008; Oskarsson et al. 2017; Persson 2015). It also quite possible that the effect of unobserved individual-level variables that are correlated with education could differ systematically between high and low QoG-countries, which would explain the conditional effect of QoG that Charron and Rothstein (2016) find.

A.9 Single-item social trust measure

The models below in Table A.13 use the classic social trust question as the dependent variable (*ppltrst* in ESS): 'Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?'. It was measured using a 11-point scale ranging from 0 ('You can't be too careful') to 10 ('Most people can be trusted').

Table A.13: Reform effect on single-item for social trust

	All reforms			Extension reforms		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.063*	0.102**	0.081	0.038	0.067	0.012
	(0.038)	(0.043)	(0.052)	(0.053)	(0.059)	(0.071)
High parental edu		0.490***	208.676		0.513***	266.110
		(0.027)	(137.165)		(0.041)	(181.523)
Ref × High par edu		-0.082*	-0.032		-0.091	0.078
		(0.042)	(0.077)		(0.060)	(0.107)
Adj. R^2	0.15	0.16	0.16	0.12	0.13	0.13
Countries	16	16	16	11	11	11
Observations	68755	64904	64904	40596	38476	38476
Flexible interaction	–	No	Yes	–	No	Yes

	Detracking reforms			Pure detracking reforms		
	(7)	(8)	(9)	(10)	(11)	(12)
Reform	0.103**	0.166***	0.181**	0.055	0.136	0.173*
	(0.049)	(0.057)	(0.071)	(0.075)	(0.083)	(0.089)
High parental edu		0.472***	177.797		0.559***	279.542
		(0.034)	(163.017)		(0.048)	(272.082)
Ref × High par edu		-0.092	-0.150		-0.029	-0.127
		(0.058)	(0.105)		(0.088)	(0.121)
Adj. R^2	0.18	0.19	0.20	0.074	0.090	0.089
Countries	11	11	11	3	3	3
Observations	29926	28141	28141	8257	7807	7807
Flexible interaction	–	No	Yes	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
Sample and controls mimic Table 3 with the following alternations. Models (1) and (2) add a full set of dummies for years of education (education FEs). Model (3) restricts the sample to respondents up to 40 years old, whereas Model (4) only includes respondents with at least one foreign-born parent or respondents belonging to an ethnic minority. Models (5) and (6) exchange the dependent variable to a measure of whether the respondent agrees that it is important to listen and try to understand people who are different from oneself, even if you disagree with these people. In Model (7) the dependent variable is instead an index for institutional trust.

A.10 Differentiating Eastern European reforms

One could suspect that the Eastern European reforms have different effects than the other reforms, considering that the older reforms were carried through during non-democratic rule and that social trust levels in general are lower in Eastern Europe. I have therefore rerun the main models separately on the reforms in Western and Eastern Europe (the flexible interaction models are excluded). This has been done for the reform effect both on educational attainment and on social trust. However, there are only three Eastern European reforms in total, two extension reforms (Hungary and Poland) and one detracking reform (Poland). Whereas the former two reforms were implemented before the transition to democracy, the latter was carried through in the late 1990s in Poland under democratic rule (see Table A.20). The small number of reforms means that we have to be cautious when studying only the Eastern European reforms. The more relevant test is to what degree the the main results hold when excluding these reforms. Since there are no pure detracking reforms from Eastern Europe, these reforms are excluded from the presentation.

The models in Table A.14 show that the Eastern European reforms had smaller effects on educational attainment than the reforms in Western Europe. The results for social trust are presented in Table A.15. The effects are actually somewhat *larger* in Eastern Europe than in Western Europe when studying all reforms. Note that the limited sample of Eastern European reforms and respondents inflates the standard errors for these models. For the extension reforms the differences between Eastern and Western European reforms are small and none of the effects are significant. If anything, the point estimates are larger for the Eastern European reforms but the estimates are also highly imprecise due to the limited sample. However, an important implication of these results is that the main result of small and insignificant effects of the extension reforms in the article (Table 3) is not a reflection of the inclusion of the Eastern reforms. Excluding the Eastern European reforms only contributes to marginally smaller effects of the extension reforms on social trust.

The lower pane of Table A.15 present the results for detracking reforms. The result for the Western reforms is effectively the same as those presented in the main Table 3 in the paper. The Polish detracking reform turns out to have a rather substantial positive effect, although the small sample means that these results should be interpreted cautiously. Its impact on the overall results is, however, negligible.

Table A.14: Reform effect on educational attainment: differentiating Eastern European reforms.

	All reforms				Extension reforms			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	0.128**	0.271***	0.080	0.139	0.173**	0.320***	0.137	0.255
	(0.063)	(0.070)	(0.114)	(0.184)	(0.073)	(0.087)	(0.128)	(0.227)
High parental edu		2.691***		2.678***		2.698***		2.786***
		(0.085)		(0.097)		(0.119)		(0.087)
Ref × High par edu		-0.198*		-0.095		-0.296*		-0.138
		(0.113)		(0.175)		(0.151)		(0.182)
Adj. R^2	0.24	0.32	0.12	0.24	0.32	0.39	0.034	0.16
Countries	14	14	2	2	9	9	2	2
Observations	61537	58328	8433	7813	33605	32099	6729	6146
Type of Reforms	West	West	East	East	West	West	East	East

	Detracking reforms			
	(9)	(10)	(11)	(12)
Reform	0.038	0.183	0.029	-0.029
	(0.107)	(0.114)	(0.212)	(0.185)
High parental edu		2.765***		2.544***
		(0.118)		(0.211)
Ref × High par edu		-0.106		-0.493*
		(0.164)		(0.258)
Adj. R^2	0.11	0.21	0.040	0.19
Countries	10	10	1	1
Observations	29684	27926	1704	1667
Type of Reforms	West	West	East	East

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

Table A.15: Reform effect on social trust: differentiating Eastern European reforms.

	All reforms				Extension reforms			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	0.061*	0.080**	0.125	0.151	0.047	0.057	0.076	0.089
	(0.032)	(0.035)	(0.083)	(0.102)	(0.045)	(0.051)	(0.102)	(0.122)
High parental edu		0.335***		0.430***		0.323***		0.415***
		(0.023)		(0.055)		(0.034)		(0.070)
Ref × High par edu		-0.033		-0.144		-0.026		-0.130
		(0.034)		(0.091)		(0.049)		(0.106)
Adj. R^2	0.194	0.205	0.0207	0.0272	0.148	0.158	0.0212	0.0268
Countries	14	14	2	2	9	9	2	2
Observations	62047	58790	8473	7858	33883	32350	6749	6170
Type of Reforms	West	West	East	East	West	West	East	East

	Detracking reforms			
	(9)	(10)	(11)	(12)
Reform	0.082**	0.113**	0.355***	0.400***
	(0.040)	(0.046)	(0.115)	(0.112)
High parental edu		0.346***		0.451***
		(0.030)		(0.102)
Ref × High par edu		-0.040		-0.158
		(0.046)		(0.147)
Adj. R^2	0.251	0.262	0.00675	0.0153
Countries	10	10	1	1
Observations	29931	28153	1724	1688
Type of Reforms	West	West	East	East

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

A.11 Excluding pure detracking reforms from detracking reforms

In the article all types of detracking reforms were collapsed to the common category *detracking reforms*, regardless of whether the reforms included an element of extended education or not. Reforms that detracked education without extending compulsory education were referred to as *pure detracking reforms*. In Table A.16 the pure detracking reforms are excluded from the broader category of detracking/mixed reforms that include detracking in connection with extended education. The same kind of models as used in Table 3 in the article are rerun for the new composition of detracking reforms.

The point estimates of the detracking reform effect are very similar to the corresponding estimates in Table 3 in the article. If anything, the estimates are slightly larger for Models (1) and (2). However, the smaller number of observations makes the estimates more imprecise and the confidence level for the average effect in Model (1) falls to the 90 per cent level. The standard errors of the more demanding full interaction model become considerably larger and the effect just falls below conventional significance levels ($P = 0.104$). However, the fact that the point estimates are similar indicate that the detracking reform effect as such is not dependent upon the categorization of pure detracking reforms. That the estimates become more imprecise for a smaller sample is expected.

Table A.16: The effect of detracking reforms on social trust: Excluding pure detracking reforms from detracking reforms.

	(1)	(2)	(3)
Reform	0.098*	0.119**	0.121
	(0.052)	(0.058)	(0.074)
High parental edu		0.320***	86.165
		(0.036)	(137.912)
Ref × High par edu		-0.044	-0.077
		(0.055)	(0.113)
Adj. R^2	0.30	0.31	0.31
Countries	11	11	11
Observations	21673	20345	20345
Flexible interaction	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. Respondents aged 25 to 80 born in the country of residence. All models include reform FEs and a quadratic general effect of birth year, as well as reform-specific linear birth year and quadratic age trends. Additional controls: dummies for gender, ethnic minority, foreign-born parents (dummies for mother, father and additional dummies for whether born outside Europe) and ESS round dummies.

A.12 Wild cluster bootstrap

In the article country-by-cohort clustered robust standard errors are used. This specification was chosen on the basis that the residuals are likely to be correlated within cohorts nested in countries since reform 'treatment' mainly is identified on this level (the exception being countries with regionally implemented reforms). Clustering the standard errors by cohort within countries also has the advantage of defining a reasonably large number of clusters, ranging from 42 clusters for the pure detracking reforms up to 384 clusters when using all reforms. I can therefore avoid issues of downward bias in the standard errors due to too few clusters.[5] Nevertheless, it might be questioned whether the residuals are independent across cohorts within countries. However, to cluster the standard errors directly on the country level with only 3 to 16 clusters, using the standard cluster-robust variance matrix estimator (CRVE), would lead to a clear risk of bias. As a robustness check I therefore use the so-called wild cluster bootstrap procedure. This procedure is recommended when the number of clusters is few and it has been shown to be one of the procedures that performs best in simulations [6]

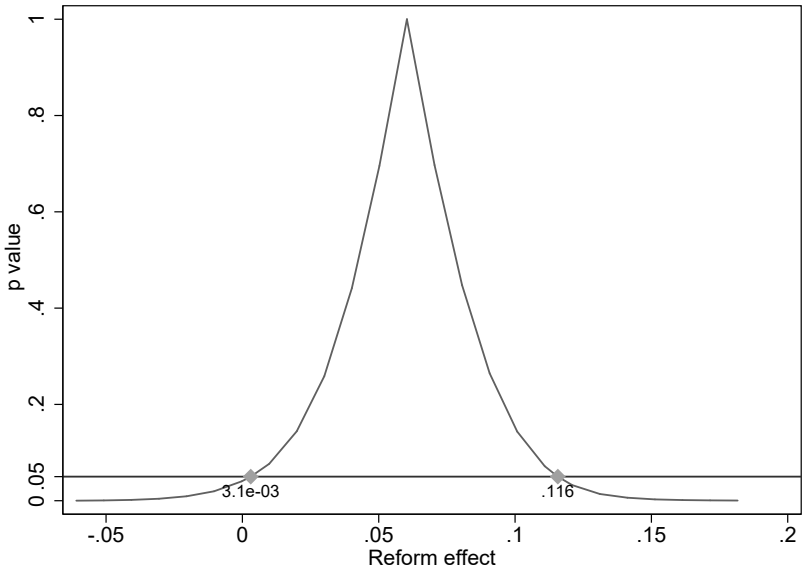
I use the *boottest* command in Stata to do these tests. Roodman et al. (2018), the developers of boottest, do not recommend using this procedure to produce standard errors but rather recommend using it to calculate P-values and so-called confidence sets (a form of confidence intervals). The reason to avoid calculating standard errors on the basis of this procedure is that such standard errors would be dependent upon a questionable normality assumption. Consequently I report p-values for testing whether some of the central coefficients in the main article are different from zero, as well as accompanying confidence sets. These confidence sets based on the bootstrap procedure (which is a non-parametric approach) differ from standard confidence intervals and are often called studentized bootstrap confidence intervals.

More formally I use the so-called WCR approach where the wild cluster bootstrap is restricted by the imposition of the null hypothesis, again in line with the recommendation of Roodman et al. (2018). Furthermore, I use the Webb six-point distribution instead of the Rademacher distribution, which is recommended when the number of clusters is small. I run 100 000 replications. In report confidence sets and p-values for the main models on all reforms, extension reforms, detracking reforms and pure detracking reforms. This is also done for the parental education interaction models but the full interaction models are excluded.

[5] Cameron, A. C. and D. L. Miller (2015). 'A practitioner's guide to cluster-robust inference'. *Journal of Human Resources* 50(2): 317-–372.

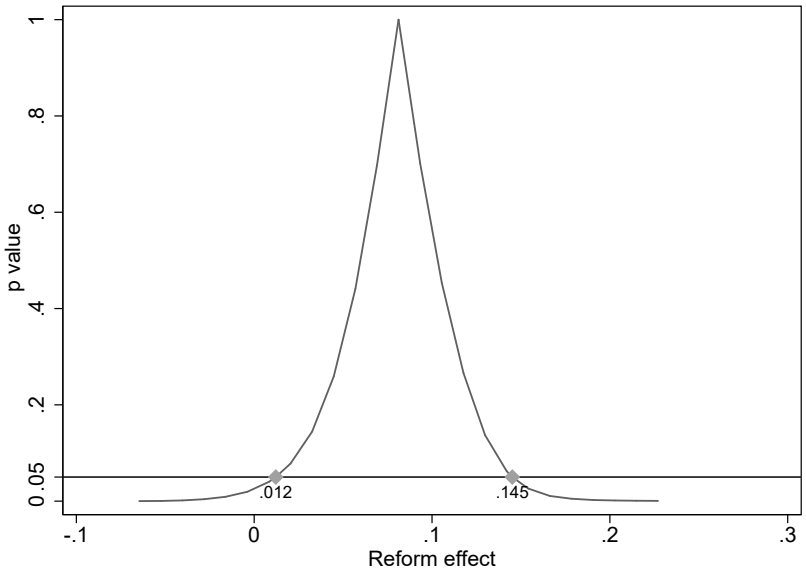
[6] See Djogbenou, A., J. MacKinnon and M.O. Nielsen . 2018, 'Asymptotic Theory and Wild Bootstrap Inference with Clustered Errors', Queen's Economics Department Working Paper No. 1399; and Roodman D., J. MacKinnon and M.O. Nielsen and M. Webb. 2018., 'Fast and Wild: Bootstrap Inference in Stata using Bboottest', Queen's Economics Department Working Paper No. 1406.)

Figure A.6: Wild cluster confidence interval for the average reform effect of all reforms. Based on Model (1) in Table 3 in the article. 95 % confidence.



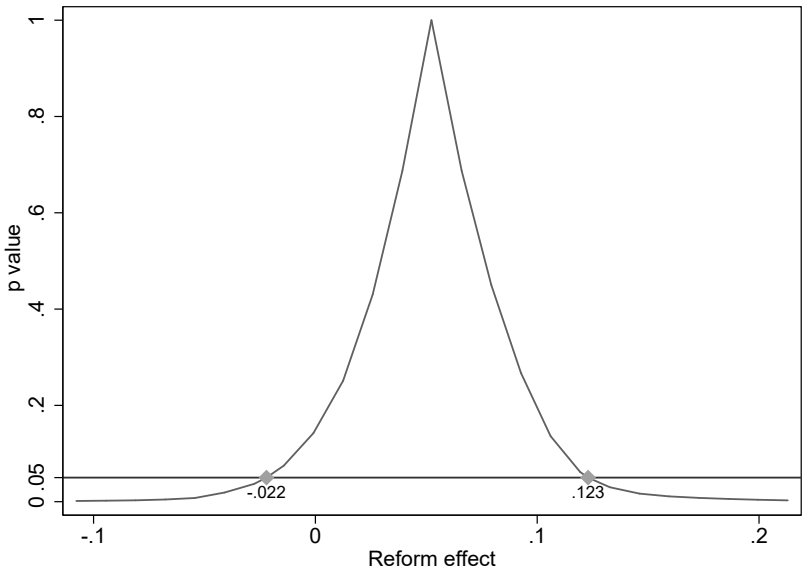
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(15) = 2.3632, Prob>|t| = 0.0409
95 per cent confidence set for null hypothesis expression: [0.003056, 0.1157].

Figure A.7: Wild cluster confidence interval for the reform effect of all reforms for individuals with poorly educated parents. Based on Model (2) in Table 3 in the article. 95 % confidence.



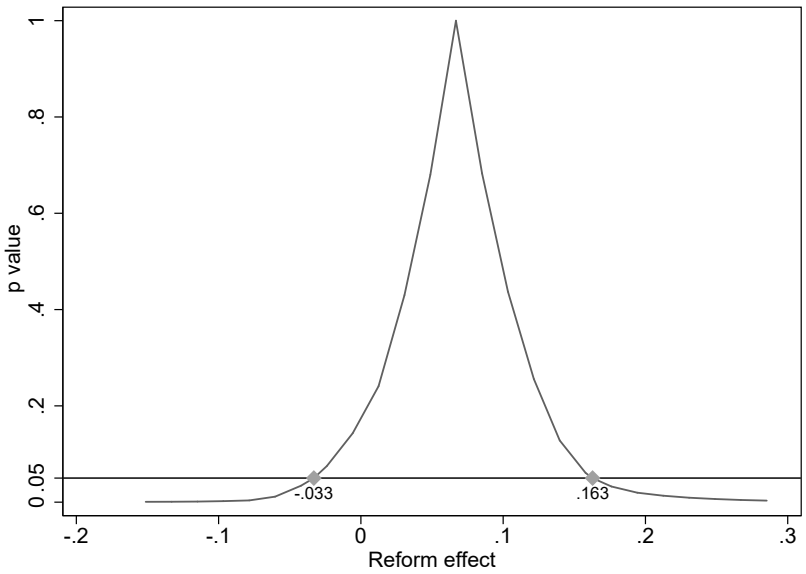
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(15) = 2.7123, Prob>|t| = 0.0246

Figure A.8: Wild cluster confidence interval for the average reform effect of extension reforms. Based on Model (4) in Table 3 in the article. 95 % confidence.



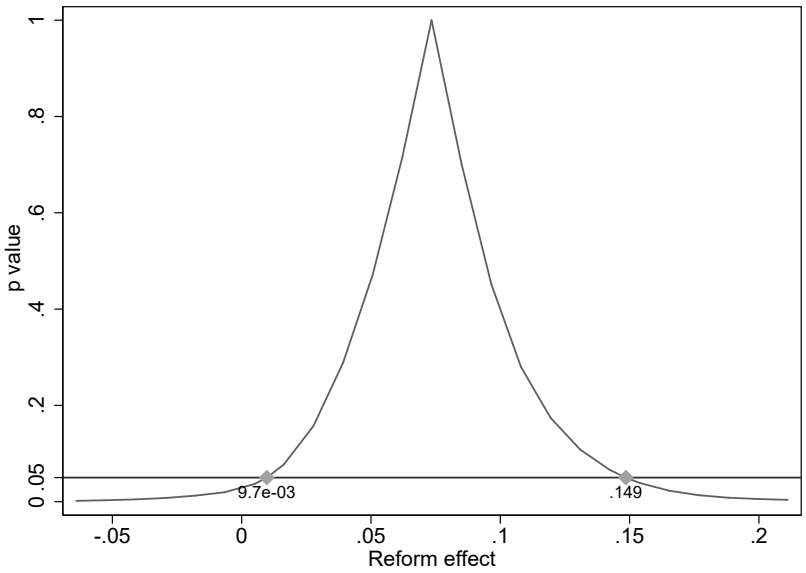
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(10) = 1.6617, Prob>|t| = 0.1480

Figure A.9: Wild cluster confidence interval for the reform effect of extension reforms for individuals with poorly educated parents. Based on Model (5) in Table 3 in the article. 95 % confidence.



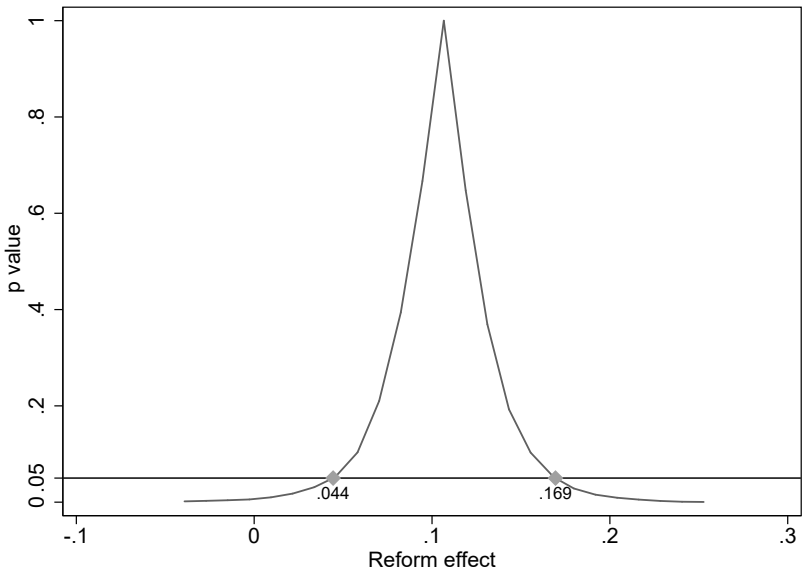
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(10) = 1.5755, Prob>|t| = 0.1688

Figure A.10: Wild cluster confidence interval for the average reform effect of detracking reforms. Based on Model (7) in Table 3 in the article. 95 % confidence.



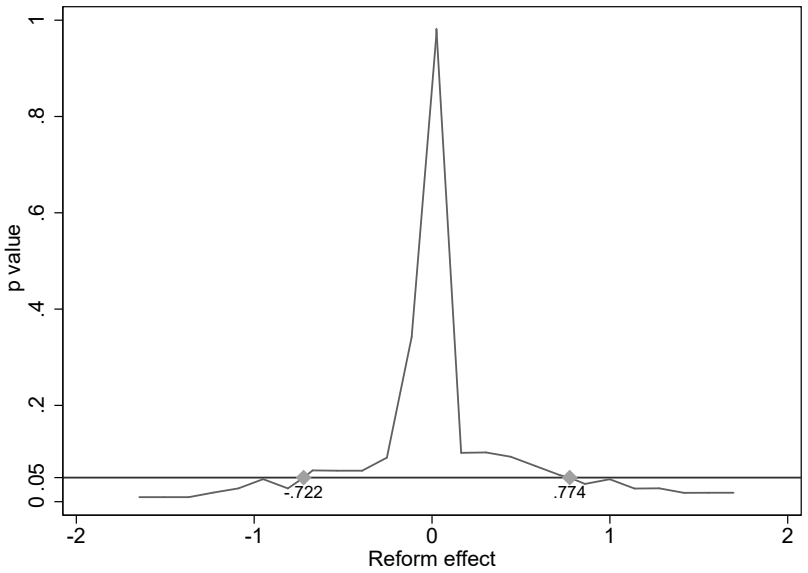
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(10) = 2.4131, Prob>|t| = 0.0270

Figure A.11: Wild cluster confidence interval for the reform effect of detracking reforms for individuals with poorly educated parents. Based on Model (8) in Table 3 in the article. 95 % confidence.



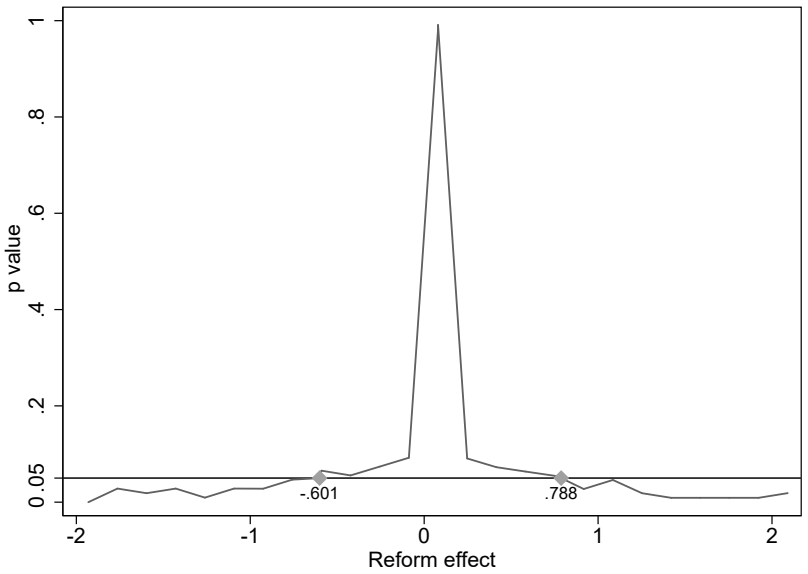
Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(10) = 3.9864, Prob>|t| = 0.0064

Figure A.12: Wild cluster confidence interval for the average reform effect of pure detracking reforms. Based on Model (10) in Table 3 in the article. 95 % confidence.



Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(2) = 0.4033, Prob>|t| = 0.6104

Figure A.13: Wild cluster confidence interval for the reform effect of pure detracking reforms for individuals with poorly educated parents. Based on Model (11) in Table 3 in the article. 95 % confidence.



Wild bootstrap, null imposed, 100 000 replications, Wald test, bootstrap clustering by country, Webb weights.
t(2) = 1.3884, Prob>|t| = 0.3781

A.13 Birth year fixed effects

In the article, the main model specification includes a general quadratic birth year trend and reform-specific linear trends for birth year and reform-specific quadratic age trends. In the models below in Table A.17 a more flexible approach to control for cohort trends is used by applying birth year fixed effects instead of the general quadratic birth year trend. These birth year fixed effects are not reform specific because they would then absorb almost all of the reform effect. This is because the reform effect is identified on the basis of birth cohorts within countries (except for the few cases with regional variation within countries in reform implementation).

It should be noted that birth year fixed effects affect reforms differently depending on the number of simultaneous reforms. Some reforms do not overlap with any other reforms in other countries and do not include any within-country variation. Birth year fixed effects then eliminate all of the variation in these reforms between the treated and non-treated birth cohorts. This mostly affects the 1947 extension reform in the UK and the 1981 detracking reform in the Netherlands. This implies an unfortunate weighting of the influence of reforms depending on the number of simultaneous reforms, which is my main reason to only using birth year fixed effects as a robustness check. The overlap of the pure detracking reforms is minimal so these models are not possible to estimate with birth year fixed effects and are therefore excluded.

The pattern is similar to the main model specification. If anything, the difference between extension and detracking reforms becomes larger as the point estimate of the effect of the extension reforms decreases somewhat in Models (4) and (5), while the point effect of the detracking reforms in Models (7) and (8) becomes slightly larger. The flexible interaction Models, (6) and (9), are more affected, which is probably explained by that interacting high parental education with a full set of birth year dummies may absorb a lot of variation. In the case of extension reforms the point estimate is zero, whereas the point estimate of the reform effect of detracking reforms is less affected but the standard error increases substantially and the effect is therefore no longer statistically significant.

Table A.17: Reform effect on social trust: birth year fixed effects (otherwise as Table 3 in the essay)

	All reforms			Extension reforms		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.040	0.057	0.031	0.047	0.052	0.000
	(0.033)	(0.036)	(0.045)	(0.048)	(0.052)	(0.063)
High parental edu		0.342***	-52.728		0.337***	-1.126
		(0.022)	(60.512)		(0.032)	(3.402)
Ref × High par edu		-0.047	0.017		-0.051	0.140
		(0.033)	(0.070)		(0.045)	(0.098)
Adj. R^2	0.20	0.21	0.21	0.17	0.18	0.18
Countries	16	16	16	11	11	11
Observations	68796	64960	64960	40632	38520	38520
Flexible interaction	–	No	Yes	–	No	Yes

	Detracking reforms		
	(7)	(8)	(9)
Reform	0.096**	0.125**	0.097
	(0.044)	(0.048)	(0.062)
High parental edu		0.349***	-2.021
		(0.030)	(3.493)
Ref × High par edu		-0.042	0.012
		(0.046)	(0.112)
Female	0.059***	0.056***	0.042
	(0.022)	(0.021)	(0.027)
Ethnic minority	-0.240**	-0.226**	-0.240**
	(0.099)	(0.103)	(0.120)
Adj. R^2	0.25	0.26	0.26
Countries	10	10	10
Observations	29931	28153	28153
Flexible interaction	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Country-by-birth cohort clustered standard errors in parentheses. The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
All models include reform FEs, a general quadratic birth year trend and reform-specific trends for birth year (linear) and age (quadratic).

A.14 Controlling for labour market status

The models below add dummy controls on the labour market status of the respondent. All of these dummies make use of the ESS items where the respondent describe his or her activities during the last 7 days. I add dummies for whether the respondent reports that he or she during the last 7 days have been in paid work, in education or retired. The reference category is the other type of activities in the ESS: unemployed and actively looking for a job, unemployed and not actively looking for a job, permanently sick or disabled, community or military service, housework and other.

These controls could potentially be so-called 'bad controls' since they may be affected by the reform (Angrist and Pischke 2009). I.e. the education reforms may, for instance, impact the likelihood of that the affected individuals become successful on the labour market, which in turn may affect social trust levels. However, I include these controls as a robustness check. There are only marginal differences in the coefficient estimates compared to the main results. This indicates that the effect on social trust of detracking reforms is not related to labour market status.

Table A.18: Reform effect on social trust: controlling for labour market position (otherwise as Table 3 in the essay)

	All reforms			Extension reforms		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.036	0.050	0.026	0.041	0.044	-0.007
	(0.033)	(0.036)	(0.045)	(0.047)	(0.052)	(0.063)
High parental edu		0.314***	-46.943		0.311***	-1.086
		(0.022)	(61.680)		(0.031)	(3.382)
Ref × High par edu		-0.044	0.017		-0.048	0.146
		(0.032)	(0.070)		(0.045)	(0.098)
Working	0.404***	0.355***	0.353***	0.369***	0.321***	0.319***
	(0.021)	(0.021)	(0.020)	(0.029)	(0.027)	(0.027)
In education	0.355***	0.306***	0.306***	0.402***	0.322***	0.321***
	(0.047)	(0.048)	(0.049)	(0.072)	(0.073)	(0.076)
Retired	0.219***	0.194***	0.188***	0.222***	0.192***	0.187***
	(0.033)	(0.033)	(0.032)	(0.040)	(0.040)	(0.039)
Adj. R^2	0.21	0.22	0.22	0.17	0.18	0.18
Countries	16	16	16	11	11	11
Observations	68796	64960	64960	40632	38520	38520
Flexible interaction	–	No	Yes	–	No	Yes

	Detracking reforms			Pure detracking reforms		
	(7)	(8)	(9)	(10)	(11)	(12)
Reform	0.081**	0.109**	0.122**	0.036	0.085	0.137*
	(0.039)	(0.044)	(0.057)	(0.051)	(0.057)	(0.073)
High parental edu		0.316***	159.256		0.372***	232.967
		(0.030)	(101.286)		(0.052)	(170.322)
Ref × High par edu		-0.038	-0.094		-0.011	-0.145
		(0.046)	(0.085)		(0.079)	(0.107)
Working	0.439***	0.394***	0.391***	0.466***	0.429***	0.425***
	(0.031)	(0.031)	(0.030)	(0.050)	(0.052)	(0.050)
In education	0.328***	0.298***	0.307***	0.300***	0.264**	0.271**
	(0.060)	(0.061)	(0.060)	(0.102)	(0.105)	(0.104)
Retired	0.186***	0.175***	0.168***	0.388***	0.349**	0.348**
	(0.057)	(0.057)	(0.057)	(0.124)	(0.143)	(0.144)
Adj. R^2	0.26	0.27	0.27	0.076	0.088	0.088
Countries	11	11	11	3	3	3
Observations	29931	28153	28153	8258	7808	7808
Flexible interaction	–	No	Yes	–	No	Yes

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Huber-White robust standard errors in parentheses.
The flexible interaction models interact all explanatory variables with parental education (the coefficient for parental education is not interpretable in these models).
All models include reform FEs, a general quadratic birth year trend and reform-specific trends for birth year (linear) and age (quadratic).

Reform overview

Table A.19: Reforms per country. Types of reforms: E, extension reform, D, pure detracking reform, ED, detracking reform (extended schooling and detracking)

Country	Region	Year implemented	First affected cohort	Type of reform	Change in comp. schooling	Note	Source
Austria		1962	1947	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006.
Belgium		1971	1960[ab]	D		After reform: 1960–64	Fort 2006; Wielemans 1991; Unesco 1971; 1977.
Belgium		1983	1969[a]	E	8 to 12	Before reform: 1965–68	Brunello et al 2009; Fort 2006
Denmark		1958	1948[a]	ED	4 to 7	Compreh. school 1–7, after r.: 1949–52.	Fort 2006.
Denmark		1971	1957[a]	E	7 to 9	Before reform: 1953–56.	Borgonovi 2010; Brunello et al 2009; Fort 2006.
Finland	Uusima[c]	1977	1966	ED	6 to 9	Compreh. school 1–9, regional impl.[d]	Borgonovi 2010; Brunello et al 2009; Fort 2006; Pekkarinen et al 2009.
Finland	Etela-Soumi[c]	1976	1965	ED	6 to 9	Compreh. school 1–9, regional impl.[d]	Borgonovi 2010; Brunello et al 2009; Fort 2006; Pekkarinen et al 2009.
Finland	Ita-Soumi[c]	1974	1963	ED	6 to 9	Compreh. school 1–9, regional impl.[d]	Borgonovi 2010; Brunello et al 2009; Fort 2006; Pekkarinen et al 2009.
Finland	Vali-Suomi[c]	1973	1962	ED	6 to 9	Compreh. school 1–9, regional impl.[d]	Borgonovi 2010; Brunello et al 2009; Fort 2006; Pekkarinen et al 2009.
Finland	Pohjois-Suomi[c]	1972	1961	ED	6 to 9	Compreh. school 1–9, regional impl.[d]	Borgonovi 2010; Brunello et al 2009; Fort 2006; Pekkarinen et al 2009.
France		1967	1953	E	8 to 10	Berthoin reform	Borgonovi 2010; Brunello et al 2009; Fort 2006.
France		1977	1966[a]	D		Before reform: 1960–65	Borgonovi 2010; Brunello et al 2009; Fort 2006.
Germany (FRG)	Schleswig-Holstein[c]	1956	1941	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Hamburg[c]	1949	1934	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Niedersachsen[c]	1962	1947	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Bremen[c]	1958	1943	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Nordrhein-Westp.[c]	1967	1953	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Hessen[c]	1967	1953	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Rheinland-Pfalz[c]	1967	1953	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Baden-Württ.[c]	1967	1953	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Bayern[c]	1969	1955	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Germany (FRG)	Saarland[c]	1964	1955	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Pischke et al 2005.
Greece		1975	1963	ED	6 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006 Gouvias 1998.
Hungary		1960	1946	E	8 to 10		Borgonovi 2010; Unesco 1962.
Ireland		1972	1958	E	8 to 9		Borgonovi 2010; Brunello et al 2009; Fort 2006.
Ireland		2000	1985	E	9 to 10		Fort 2006; UNESCO 2001.
Italy		1963	1950	ED	5 to 8	Compreh. school 1–8	Flabbi 1999; Fort 2006.

Table A.20: Reforms per country, continued

Country	Region	Year implemented	First affected cohort	Type of reform	Change in comp. schooling	Note	Source
The Netherlands		1968/75	1958[e]	ED	9 to 10	Mammoth law & extended comp. schooling	Fort 2006; Levin et al 1999; Oosterbeek et al 2004; Luijkx et al 2008.
The Netherlands		1993	1981[a]	D		After reform: 1981–83	Fort 2006; Levin et al 1999.
The Netherlands		1998	1986[a]	D		Before reform: 1984-85	Luijkx & de Heus 2008.
Poland		1961	1952	E	7 to 8	Extended primary schooling[f]	Borgonovi 2010; Szebenyi 1992.
Poland		1999	1986	ED	8 to 9	Compreh. school 1–9	Jakubowski et al. 2010.
Portugal		1956	1948[a]	E	3 to 4	Only boys, after reform: 1948–51	Fort 2006; Vieira 1999.
Portugal		1960	1952[a]	E	3 to 4	Only girls, after reform: 1952–53[g]	Fort 2006; Vieira 1999.
Portugal		1964	1956[a]	E	4 to 6	Before reform: boys 1952–55, girls 1954–55	Fort 2006.
Portugal		1987	1981	ED	6 to 9		Fort 2006.
Spain		1970	1957	ED	6 to 8		Fort 2006; Borgonovi 2010.
Spain		1991	1977	E	8 to 10	LOGSE and increased comp. education[h]	Zamarro 2010, Felgueroso et al 2014.
Sweden		1962	1951	ED	7/8 to 9[i]	Compreh. school 1–9, regional impl.[j]	Borgonovi 2010.
UK	England[c]	1947	1933	E	8 to 9		Fort 2006; Tomlinsson 1991; Oreopoulos 2006.
UK	England[c]	1972	1957	E	9 to 10	Gradual impl. of comprehensive schools[k]	Fort 2006; Borgonovi 2010; Grenet 2013.
UK	Wales[c]	1947	1933	E	8 to 9		Fort 2006; Tomlinsson 1991; Oreopoulos 2006.
UK	Wales[c]	1972	1957	E	9 to 10	Gradual impl. of comprehensive schools[k]	Fort 2006; Grenet 2013.
UK	Northern Ireland[c]	1957	1943[a]	E	8 to 9	Before reform 1941–42, after reform: 1944–49	Fort 2006; Oreopoulos 2006
UK	Northern Ireland[c]	1972	1957	E	9 to 10	Before reform: 1950–56	Fort 2006; Borgonovi 2010.
UK	Scotland[c]	1947	1933	E	8 to 9		Fort 2006; Oreopoulos 2006.
UK	Scotland[c]	1972	1957	E	9 to 10		Fort 2006.

Comments and notes

The focus on reforms that only extend or detrack education, and not any reforms in the opposite direction may spur questions. However, reforms that have shortened compulsory schooling or increased the level of tracking, have been rare. These have usually been implemented during extraordinary circumstances such as under military dictatorship (Portugal 1926, Greece 1967–1974) or when several of the former Socialist republics in Eastern Europe reintroduced tracked educational systems after democratization (e.g., The Czech Republic and East-Germany). The social context of these reforms make them unsuitable to exploit.

[a] For these reforms the standard $+/-7$ year window before and after reform would lead to overlap with another reform. The reform window is adjusted according to the Note column.

[b] The reform targeted public secondary schools which the majority of children attended. Independent schools, mostly religion-based catholic schools, could continue with the traditional system but this also varied between regions. See Wielemans (1991).

[c] I only have data for where the respondents lived when responding to the survey, not where they went to school. I consequently have to assume that the respondents still live in the same region/state where they went to school. For the German case there is good support for this assumption. Using data from the German ALLBUS-surveys, Pischke (2003) finds that 84 % of the respondents during the 1990s, still lived in the state were they born.

[d] The reform was gradually implemented on a municipal level. As I lack data for the municipality of the respondents, the reform implementation is approximated on a regional level according to the year in which the largest share of municipalities in the region implemented the new comprehensive school. This follows the procedure of Borgonovi et al. (2010) and Brunello et al. (2009).

[e] A reform ('Mammoth law') implemented in 1968 made secondary school less tracked, affecting cohorts from 1957. A reform implemented in 1975 extended compulsory schooling, affecting cohorts from 1959. These reforms are treated as one reform implying both extended comp. schooling and detracking, affecting cohorts born 1958 and later.

[f] This reform extended compulsory primary schooling from seven to eight years in Poland. However, the secondary system in which students could choose between different vocational, technical and academic schools was unchanged. As the reform did not affect secondary schooling the reform is coded as an extension reform and not a detracking reform.

[g] For this reform the first affected cohort is included to avoid a too small sample of treated individuals.

[h] This reform actually consists of two different reforms. First, compulsory schooling was increased from age 14 to 16, affecting everyone from 1991. Second, a new non-tracked compulsory secondary school for 12 to 16 year olds was introduced, LOGSE (*Ley Orgánica de Ordenación General del Sistema Educativo*). The LOGSE was implemented gradually across regions. However, while the increase in compulsory schooling was implemented immediately 1991, the LOGSE was implemented gradually through the 1990s and affected only a small share of each cohort during the first years (less than 10 percent of each cohort). In 1994 at most 20 percent of each cohort were affected, and it was not until 1996 about 50 percent of each cohort studied under the new system (see Felgueroso et al, 2014). Unfortunately I am unable to make use of the regional variation due to data limitations. Since the introduction of the LOGSE did not really coincide with the increase in compulsory schooling (more than for a very small share of the sample) and did not affect more than a small share of each cohort for most of the seven years following 1991, I do not classify this reform as a detracking reform. Most of the respondents in the ESS belonging to the cohorts that to a larger extent was affected by the LOGSE would also fall below my lower age limit of 25 years old.

[i] Before the reform compulsory schooling differed between 7 and 8 years, depending on the local municipality.

[j] This reform was implemented gradually on a municipal level. However, due to that I am unable to map the survey respondents to the the municipalities I cannot make use of this variation. The first affected cohort is instead the cohort for which most municipalities implemented the new comprehensive school. I.e. the year with the largest shift (cf. Borgonovi et al. 2010).

[k] This reform implied an immediate extension of compulsory schooling. However, during the 1960s and 1970s a gradual implementation of comprehensive schools was also ongoing, but implementation was decentralized to the local education authorities. Due to data limitations I may not exploit this gradual detracking of the English and Welsh education systems. Since the implementation was continuously ongoing at the time, picking up speed already 1965 with the government circular, N. 10/65, this gradual development should have a limited, if any, confounding effect on the results of the extension reform. That is, the implementation does not coincide with the timing of the extension of compulsory schooling in 1972, as most schools at that time already had adopted the comprehensive system (Grenet 2013).

References for reform sources

Borgonovi, F., d'Hombres, B., & Hoskins, B. (2010). 'Voter turnout, information acquisition and education: evidence from 15 european countries'. *The BE Journal of Economic Analysis & Policy*, 10(1).

Brunello, G., Fort, M., & Weber, G. (2009). 'Changes in compulsory schooling, education and the distribution of wages in europe*'. *The Economic Journal*, 119(536), pp. 516—539.

Felgueroso, F., Gutiérrez-Domènech, M., & Jiménez-Martín, S. (2014). 'Dropout trends and educational reforms: the role of the LOGSE in Spain'. *IZA Journal of Labor Policy*, 3(1), pp. 1–24.

Flabbi, L. (1999). 'Returns to schooling in Italy: OLS, IV and gender differences'. Working paper n. 1, 1999, Universitá Bocconi, Serie di econometria ed economia applicata.

Fort, M. (2006). 'Educational reforms across Europe: a toolbox for empirical research'. Unpublished paper.

Gouvias, D. (1998). 'Comparative Issues of selection in Europe: the case of Greece'. *Education Policy Analysis Archives*, 6(4).

Grenet, J. (2013). 'Is extending compulsory schooling alone enough to raise earnings? Evidence from french and british compulsory schooling laws*'. *The Scandinavian Journal of Economics*, 115(1), pp. 176–210

Jakubowski, M., H. A. Patrinos, E. E. Porta, J. Wisniewski (2010). The impact of the 1999 Education Reform in Poland

Levin, J., & Plug, E. J. (1999). 'Instrumenting education and the returns to schooling in the Netherlands'. *Labour Economics*, 6(4), pp. 521–534.

Luijkx, R., & de Heus, M. (2008). 'The educational system of the Netherlands'. *The international standard classification of education (ISCED-97). An evaluation of content and criterion validity for 15 European Countries*, pp. 47-75.

Oosterbeek, H. & Webbink, D. (2004). 'Wage effects of an extra year of lower vocational education: evidence from a simultaneous change of compulsory school leaving age and program length'. Scholar Working Paper Series; No. WP 44/04, Amsterdam: University of Amsterdam.

Oreopoulos, P. (2006). 'Estimating average and local average treatment effects of education when compulsory schooling laws really matter'. *The American Economic Review* 96.1, pp. 152–175.

Pekkarinen, T., Uusitalo, R. & Kerr, S. (2009). 'School Tracking and Development of Cognitive Skills'. IZA Discussion Paper 4058.

Pischke, J.-S. (2003). 'The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Year'. NBER Working Paper 9964.

Pischke, J.-S. & Von Wachter, T. (2005). 'The impact of length of the school year on student performance and earnings: evidence from the german short school year'. NBER Working Paper 11414.

Szebenyi, P. (1992). 'Change in the systems of public education in east central Europe'. *Comparative Education*, 28(1), pp. 19–31

Tomlinson, J. R. G. (1991). 'Comprehensive education in England and Wales', 1944-1991. *European Journal of Education*, pp. 103-117.

UNESCO (1962). *International yearbook of education, vol. XXIII, 1961.* International Bureau of Education and UNESCO.

UNESCO (1971). *Some views on the educational evolution in 1970-1971.* Belgique – Ministries of National Education and Culture. Prepared for the XXXIIIrd International Conference on Education.

UNESCO (1977). *Le mouvement éducatif en Belgiue 1974-1976.* Belgium – Ministère de l'education nationale. 36th Session of the International Conference on Education.

UNESCO (2001). *National report of Ireland.* International Conference on Education, 46th Session, Geneva 2001.

Vieira, J. A. (1999). 'Returns to education in Portugal'. *Labour Economics*, 6(4), pp. 535-541.

Wielemans, W. (1991). 'Comprehensive education in Belgium: a broken lever?' *European Journal of Education*, 26(2), pp. 167–178.

Zamarro, G. (2010). 'Accounting for heterogeneous returns in sequential schooling decisions'. *Journal of Econometrics*, 156(2), pp. 260-276.