Online appendix for "Parents, Peers, and Politics: The Long-Term Effects of Vertical Social Ties"

Linuz Aggeborn Nazita Lajevardi Karl-Oskar Lindgren
Pär Nyman Sven Oskarsson

Details on data and measures

This section provides a description of the data availability, data sources and the main variables used for the paper "Parents, Peers, and Politics: The Long-Term Effects of Vertical Social Ties."

Data availability and replication

We use individual level data from Swedish registers. The data material is located on an encrypted server to which we have to log in through a remote desktop application in order to perform all of our data analyses. Due to the extreme sensitivity of the data, we are under contractual and ethical obligation not to distribute these data to others. For

*The authors are affiliated with Department of Government, Uppsala University (LA, KOL, PN and SO), Uppsala Center for Labor Studies (LA, KOL, PN, SO), Michigan State University (NL), Uppsala Center for Fiscal Studies (LA) and Institute for Evaluation of Labor Market and Education Policy (KOL). Corresponding author's address: Linuz Aggeborn, Box 514, 75120 Uppsala, Sweden. E-mail: linuz.aggeborn@statsvet.uu.se. We thank seminar participants at Uppsala University, the Oslo Turnout Workshop (2017), Swepsa (2017), APSA (2018), and MPSA (2019) for their helpful and constructive comments. Earlier versions of this paper circulated under the titles "Do Political Acquintances Make You Politically Active?" and "The Effect of Having Peers Whose Parents Are Politicians". This research was funded by the European Research Council (ERC), grant number 683214 CONPOL, and the Swedish Research Council (VR), grant number 2017-02472.

that reason, we requested an exception from the journal's data and replication policy at the time of first submission. The editors granted us such an exception.

For those researchers who want to replicate our results there are two ways to get access to the administrative data. The first way is to order the data directly from Statistics Sweden (SCB). Statistics Sweden presently requires that researchers obtain a permission from a Swedish Ethical Review Board before data can be ordered (a description, in Swedish, of how to order data from Statistics Sweden is available at: `https://www.scb.se/en/services/guidance-for-researchers-and-universities/`). We will also make available a complete list all of the variables that we ordered from Statistics Sweden for this project, together with the dofiles and Stata logs.

The second way to replicate our analyses is to come to Sweden and reanalyze these data through the same remote server system that we used. Researchers interested in using this option should reach out to us prior to coming to Sweden so that we can apply for approval from the Ethical Review Board for the researcher to temporarily be added to our research team, which is mandatory in order to get access to the remote server system.

Variables and data sources

Voter turnout

The Swedish registers do not contain population-wide turnout information. Although Statistics Sweden (SCB) has collected information on individual turnout for each election since 1991, their samples only cover about 1 percent of the electorate. However, the electoral rolls are still maintained in paper form, and each roll lists all eligible voters living a particular voting district. The electoral rolls contain preprinted information on the full name and a unique personal identification number (*personnummer*) for all eligible voters, and hand-written information, filled in by the election officials, on whether particular individuals chose to vote in each of the three different elections at the municipal, county and national levels. By scanning and digitizing these election rolls, population data on voter turnout in the 2009 European parliament election and the 2010 general election ($N \approx 7,000,000$) could be collected. Comparisons show that the data conforms with

the data collected by Statistics Sweden in 99.7 percent of the cases (85,235/85,449). See Lindgren et al. (2017) for a description of the procedures with regards to to scanning and digitizing these election rolls.

Data from administrative registers

In the main analysis we make use of data from various administrative registers. In this subsection we describe the main variables in somewhat more detail.

`School class` – The unique combination of school, program, and year of application to upper secondary school. The information is retrieved from the Upper Secondary School Application Record (*Gymnasieskolans sökanderegister*).

`Number of politicians` – The number of parents to the children in a class, divided by the number of children and multiplied by 25 for comparison between classes of different sizes. We choose 25 students because it should represent a fairly standard class in upper secondary school. A politician is defined as a parent who was nominated either in the election prior to, or the election during, when the child was in upper secondary school. The data on politicians comes from the Register of Candidates and Elected. We use the multigenerational data set to connect the individuals to their parents in the data set.

`Ever nominated` – A binary variable for whether the person was a running (i.e. being nominated) in at least one of the municipal, regional or national elections between 1982 and 2014 (parents) or 1998 and 2014 (children). The data for 1991–2014 come from the Register of Candidates and Elected, whereas the data for the years 1982–1988 have been gathered by Olle Folke and Johanna Rickne.

`Ever elected` – A binary variable for whether the person was elected in at least one of the municipal, regional or national elections between 1982 and 2014 (parents) or 1998 and 2014 (children). The data comes from the Register of Candidates and Elected.

`Voter turnout, European election` – A binary variable whether the individual, father or mother voted in the European Election in 2009

`Voter turnout, general election` – A binary variable whether the individual, father or mother voted in the parliamentary election in 2010

`Mother nominated after` – A binary variable for whether the mother was running (i.e. being nominated) in at least one of the municipal, regional or national elections after the child graduated from upper secondary school. The data comes from the Register of Candidates and Elected.

`Father nominated after` – A binary variable for whether the father was running (i.e. being nominated) in at least one of the municipal, regional or national elections after the child graduated from upper secondary school.. The data comes from the Register of Candidates and Elected.

`Political discussion` – Four-category indicator measuring whether the respondent engages in political discussions. This measure is based on the following survey item:

> How do you behave when you are in a group and political questions are discussed?
> (1) I don't listen when people talk politics
> (2) I usually listen, but I never participate in the discussion
> (3) I sometimes express my opinions
> (4) I usually participate in the discussion and voice my opinions

The information is retrieved from the annual Living Conditions Surveys 1997–2014 (ULF by Swedish acronym). The sample size in these surveys is around 10,000 individuals and the response rate averaged approximately 65-70%.

`Municipality of residence` – Code for the municipality of residence to be used as municipality fixed effects. The information originates from the 2009 wave of the Longitudinal integration database for health insurance and labour market studies (LISA).

`Gender` – Equal to 1 if female and 0 for male. The information originates from the Swedish Population Register.

`Foreign born` – Equal to 1 if the individual is foreign born and 0 if born in Sweden. The

information comes from the Swedish Population Register.

`Years of education, father` – Father's highest education, expressed in years. The education levels are based on the Swedish standard classification of education (SUN 2000). The years of education are set to 6.6/7.5/9.4/11.2/12.4/14.2/17.0/20.4 if the highest SUN-level is 1/2/3/4/5/6/7. If the SUN-level is missing, "years of education" is set to 6.6. The information is originates from the LISA database and is measured during the same year as the child graduates from upper secondary school.

`Years of education, mother` – Mother's highest education, in years. The education levels are based on the Swedish standard classification of education (SUN 2000). The years of education are set to 6.6/7.5/9.4/11.2/12.4/14.2/17.0/20.4 if the highest SUN-level is 1/2/3/4/5/6/7. If the SUN-level is missing, years of education are set to 6.6. The information is originates from the LISA database and is measured during the same year as the child graduates from upper secondary school.

`Standardized income, father` – Father's gross wage, standardized within each year to a variable with the mean 0 and standard deviation 1. The information originates from the LISA database.

`Standardized income, mother` – Mother's gross wage, standardized within each year to a variable with the mean 0 and standard deviation 1. The information is retrieved from the LISA database.

`Social assistance recipient, father` – A binary variable for whether the father lived in a family that received social assistance during the child's graduation year. The information originates from the LISA database.

`Social assistance recipient, mother` – A binary indicator for whether the mother lived in a family that received social assistance during the child's graduation year. The information originates from the LISA database.

`Employed, father` – A binary variable for whether the father was employed during the year the child graduates from upper secondary school. The information comes from the

LISA database.

`Employed, mother` – A binary variable for whether the mother was employed during the year the child graduates from upper secondary school. The information comes from the LISA database.

`Cognitive abilty` – A variable standardized with mean 0 and standard deviation 1 for each graduation year. This variable is mostly available for men and originates from the enlistment data base.

`Non-cognitive abilty` – A variable standardized with mean 0 and standard deviation 1 for each graduation year which measure social skills. Non-cognitive ability was assessed by a psychologist during conscription. This variable is mostly available for men and originates from the enlistment data base.

`GPA elementary school` – A variable standardized with mean 0 and standard deviation 1 for each graduation year and grade system. The variable originates from the register for ninth grade from Statistics Sweden.

`GPA upper secondary` – A variable standardized with mean 0 and standard deviation 1 for each graduation year and grade system. The variable originates from the Upper Secondary School Graduation Record.

Robustness Analyses

As noted in the main text, our identification strategy hinges on a number of assumptions. Here, we examine these assumptions and conduct sensitivity analyses to examine the robustness of our results to departures from them.

Trends and Dynamics

In Table 5 of the main text we present results for models with separate time trends for each combination of school and program. In Tables A1 and A2, we instead include separate school and program time trends. We further control for the number of politician parents in the classes preceding and succeeding the treated classes in question (the lead and lag share). Column 2 in Tables A1 and A2 replicate the results from Table 5 to facilitate comparison. The pattern of results is very similar to the one found in Table 5 when the trends are added separately. The estimated coefficients decreases somewhat in magnitude for voter turnout in the European election and for the two measures for elite participation but increases with regards to voter turnout in the general election when the lead and lag shares are included together with the school-program trend. Taken all together, we remain at our conclusion that the main results in Table 2 and Table 3 in the main text are robust to alternative ways of specifying the time trends and the inclusion of the number of politician parents in the lag and lead school classes.

Table A1: Voter turnout – adding trends and dynamics

	(1) Vote09	(2) Vote09	(3) Vote09	(4) Vote10	(5) Vote10	(6) Vote10
Number of politicians	0.265	0.262	0.198	0.089	0.072	0.118
	(0.070)	(0.074)	(0.087)	(0.053)	(0.055)	(0.064)
Mean dep.var.	40.815	40.815	40.638	81.256	81.256	81.099
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	Yes	Yes	Yes	Yes	Yes	Yes
School trend	Yes	No	No	Yes	No	No
Program trend	Yes	No	No	Yes	No	No
School-program trend	No	Yes	Yes	No	Yes	Yes
Lagged and lead share	No	No	Yes	No	No	Yes
Adjusted R2	0.118	0.115	0.115	0.135	0.133	0.127
Observations	909,656	909,690	764,793	1,091,302	1,091,336	936,860

Note: Results from models including linear time trends and the lagged and lead value of the treatment variable. The outcome in columns 1–3 is turnout in the 2009 EP election whereas estimates for turnout in the 2010 national election are presented in columns 4–6. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A2: Nominated and elected – adding trends and dynamics

	(1) Nom	(2) Nom	(3) Nom	(4) Elec	(5) Elec	(6) Elec
Number of politicians	0.020	0.025	0.019	0.001	0.003	0.001
	(0.012)	(0.013)	(0.014)	(0.005)	(0.006)	(0.006)
Mean dep.var.	0.615	0.615	0.573	0.114	0.114	0.104
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	Yes	Yes	Yes	Yes	Yes	Yes
School trend	Yes	No	No	Yes	No	No
Program trend	Yes	No	No	Yes	No	No
School-program trend	No	Yes	Yes	No	Yes	Yes
Lagged and lead share	No	No	Yes	No	No	Yes
Adjusted R2	0.004	0.001	0.000	0.001	−0.002	−0.002
Observations	1,099,105	1,099,139	943,458	1,099,105	1,099,139	943,458

Note: Results from models including linear time trends and the lagged and the lead value of the treatment variable of the number of politicians in a class. The outcome in columns 1–3 is running for office at least once in the five elections held between 1998 and 2010 whereas columns 4–6 instead display results for winning office at least once in the same elections. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Sensitivity to other functional forms

In the main text in Table 2 and Table 3, we calculate the number of politician parents in a class of 25 students and use that as the independent variable in a linear specification. In this section, we investigate the effect using other functional forms. First, we define a dummy variable taking the value 1 if there is at least one politician parent in the class and 0 otherwise. Second, we run an analysis where we create five dummy variables for the quintiles in the distribution of politician parents, where the first quintile dummy is left out as a reference point. Lastly, we use the log of the number of politician parents. Given that the log of 0 is undefined we have added 1 to all values before taking the logarithm. The results are presented in Tables A3–A4.

To summarize the results in this section, we find that for more common political tasks, such as voting in the national election, the effect of having a politician parent in the class is foremost driven by having at least one politician. For voting in the European election, there seem to be a possibly linear increase of having additional politicians parents on voter turnout. For being nominated in the future, the effect is foremost driven by the fifth quintile dummy, which corresponds to a group where there on average is two politician parents per class.

Let us now discuss the outcomes one at the time. For turnout in the national elections, the binary indicator receives a coefficient that is twice the size of the coefficient for our continuous measure in the main analysis. The coefficients for quintiles 2–5 are all positive, but not increasing, indicating that there is a positive effect from the first politicians but that this effect cannot be extrapolated when the numbers of politicians increase. The interpretation for the log specification is that if the number of politician parents in a class of 25 students is increased by one percent, the probability that the individual will vote in the national election in 2010 will increase by 0.0020 percentage points.

For turnout in the European elections, the effect appears to be more linear. When interpreting the quintile coefficients, it is important to note that the classes in the first three quintile have less than one parent politician per class (in the second and third quintile, the classes are larger than 25 students). So although the only positive coefficients

are found for the fourth and fifth quintile, it is also here that we see a major difference in politicians per class (1 and 2, respectively, compared to the reference category of zero politicians). For a perfect linear relationship, the coefficients for the second and third quintile should had been 0.1 and 0.2, which they clearly are not, but these differences are within the margin of error. For the log specification, the probability of voting in the 2009 election will increase by 0.0067 percentage points if the number of politician parents is increased by 1 percent.

For elite participation, we conclude that the effect found in the main text is driven by observations where there are multiple politicians among the parents in a class. In our main specification, these observations have tons of leverage, but when we group the data into two categories (zero or more politicians) or quintiles, this leverage is lost and no group is no longer statistically significantly different from the reference category. If we instead of quintiles use ventiles, with each group except the first one spanning 5 percent of the distribution, it is clear that it is the top five percent that drives the main results (centiles show a similar pattern). In this group, the average number of politicians is 3.4 per class. Although some may be inclined to dismiss such results as driven by individual outliers, we want to emphasize that each ventile consists of more than 50,000 observations.

Table A3: Voter turnout – Different functional forms

	(1) Vote09	(2) Vote09	(3) Vote09	(4) Vote10	(5) Vote10	(6) Vote10
At least one politician	0.385			0.253		
	(0.144)			(0.107)		
Second quintile		−0.141			0.351	
		(0.272)			(0.169)	
Third quintile		0.006			0.266	
		(0.186)			(0.131)	
Fourth quintile		0.403			0.126	
		(0.177)			(0.131)	
Fifth quintile		0.835			0.321	
		(0.183)			(0.134)	
Log. number of politicians			0.671			0.201
			(0.118)			(0.081)
Mean dep.var.	40.815	40.815	40.815	81.256	81.256	81.256
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	0.118	0.118	0.118	0.135	0.135	0.135
Observations	909,690	909,690	909,690	1,091,336	1,091,336	1,091,336

Note: Results from OLS regressions. The outcome in columns 1–3 is turnout in the 2009 EP election whereas estimates for turnout in the 2010 national election are presented in columns 4–6. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A4: Elite political participation – Different functional forms

	(1) Nom	(2) Nom	(3) Nom	(4) Elec	(5) Elec	(6) Elec
At least one politician	−0.014			−0.008		
	(0.019)			(0.008)		
Second quintile		−0.052			−0.013	
		(0.034)			(0.013)	
Third quintile		−0.029			−0.013	
		(0.025)			(0.011)	
Fourth quintile		−0.023			−0.005	
		(0.026)			(0.011)	
Fifth quintile		0.025			−0.002	
		(0.029)			(0.012)	
Log. number of politicians			0.006			−0.000
			(0.018)			(0.007)
Mean dep.var.	0.615	0.615	0.615	0.114	0.114	0.114
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	0.004	0.004	0.004	0.001	0.001	0.001
Observations	1,099,139	1,099,139	1,099,139	1,099,139	1,099,139	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is turnout in the 2009 EP election whereas estimates for turnout in the 2010 national election are presented in columns 4–6. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Marginal Effects from Non-Linear Models

There are many reasons for why we prefer to use linear probability models in our main specifications in Tables 2–3 in the main text. Here we test whether we would get similar results if we used a non-linear model instead.

Our identification strategy requires that we cancel out the fixed effects for the school-program combinations. The standard approach of doing that in a non-linear model is to use conditional logit with the school-program as the grouping variable. However, for computational reasons, we cannot do that while simultaneously maintaining the same model specifications (or rather their equivalent in logistic regressions) as we used in our OLS regressions. Conditional logit relies on evaluating the binomial coefficient to calculate the number of possible ways that the observed number of voters (for example) could occur among all the individuals in a school-program combination (our grouping variable). Evaluating the binomial coefficient sometimes results in larger numbers than what is able to be represented in double precision. We have used the approach suggested by Stammann et al. (2016), which is computationally efficient and equivalent to a standard logit estimator with a dummy variable for each school-program combination.

The marginal coefficient estimates from these conditional logit models are reported in Table A5. The main take-home point here is that the pattern of results are very similar to the ones reported in the main text in Table 2 and Table 3. Above all, vertical social ties are positively related to all four measures of political participation.

Table A5: Conditional logit estimates

	(1) Voted09	(2) Voted10	(3) Nominated	(4) Elected
Number of politicians	0.283***	0.093*	0.018	0.002
	(0.068)	(0.05)	(0.023)	(0.006)
Individual covariates	Yes	Yes	Yes	Yes
Parent covariates	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Muni. FE	No	No	No	No
Lag and lead of treatment	No	No	No	No
Observations	931915	1113689	886141	474662

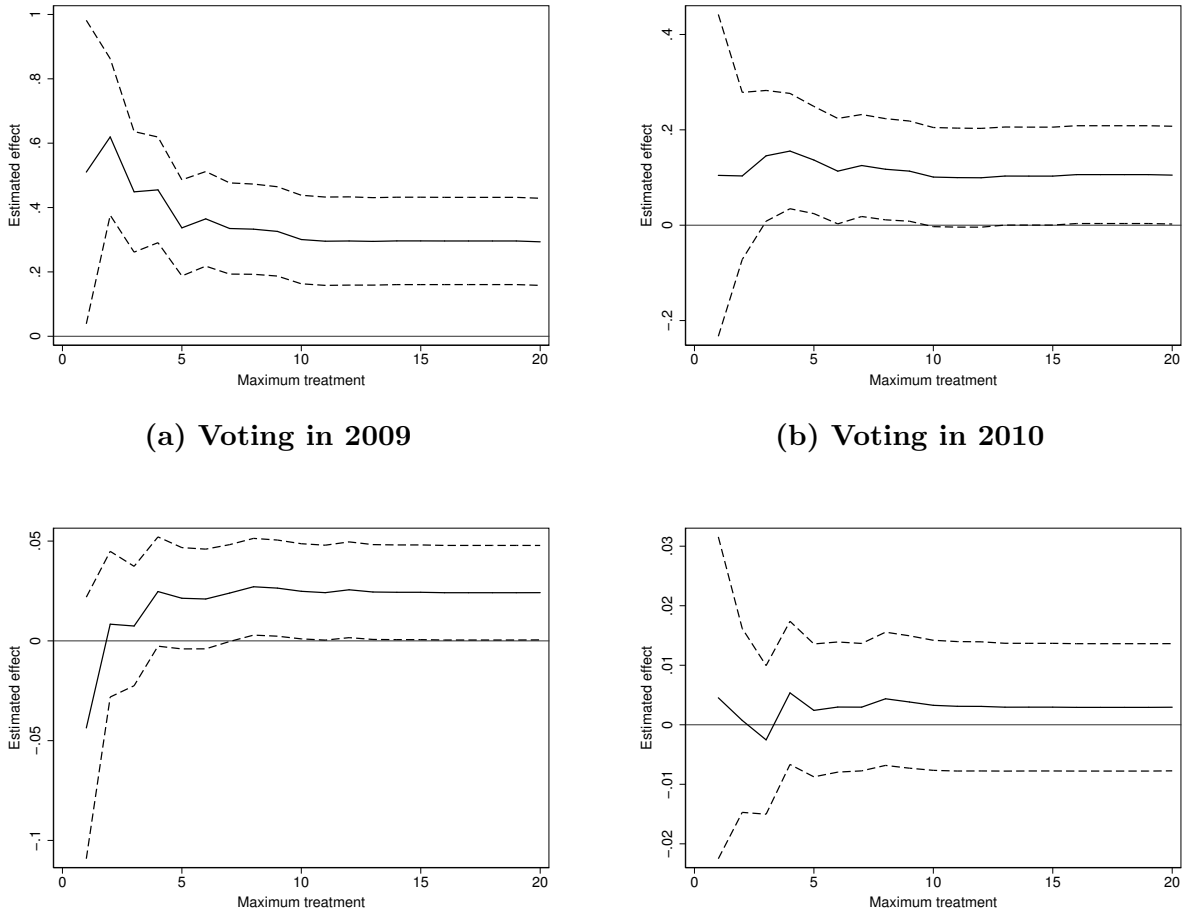
Note: Results from conditional logit models. The models include the same covariates as we used in the second (and fifth) column in Table 2 and 3 in the main text. Column 1 presents results using turnout in the 2009 EP election as outcome; in column 2 turnout in the 2010 national election is used as outcome; columns 3 and for employ running for and winning office at least once in the five elections held between 1998 and 2010. Standard errors, shown in parentheses, allow for clustering at the school-program level. Results are presented as marginal effects over the response surface (in percentage points).

Excluding Outliers

As discussed in the main text, the distribution of the treatment variable is positively skewed and includes some values that are more than 20 times larger than the average number of politicians per class. One possible cause for concern is that such outliers may unduly affect the coefficient estimates. Although outliers are a smaller problem in studies like ours, just because of the sheer number of observations, it still important to check how sensitive the results are to the inclusion of the potentially excessively influential observations. For this reason, Figure A1 shows how the estimated coefficient changes if we successively exclude observations based on the number of politicians in a class (decreasing the maximum number of politician parents per 25 students from 20 to 1). All regressions are based on our preferred specification (the third and sixth columns in Tables 2 and 3) used in the main text.

As should be expected the estimates get more noisy as we exclude school classes with a large number of politician parents from the sample. More importantly, however, there is no consistent pattern in how the point estimates change when excluding all treated classes except for those in which one or just a few of the parents run for office (the leftmost

Figure A1: Estimated effects with sample restrictions on number of politicians



(a) Voting in 2009 **(b) Voting in 2010**

(c) Ever nominated **(d) Ever elected**

Note: The graphs display coefficient estimates (solid line) and 95% confidence intervals (dashed lines) from models in which the sample is restricted such that school classes surpassing the specified maximum number of politician parents on the x-axis (between one and twenty) are excluded.

estimate in each subgraph): for turnout in the 2009 EP election the effect increases, for being nominated at least once the effect decreases, and for the two remaining outcomes (turnout in the 2010 national election and winning political office) the effect of vertical ties is stable when restricting the maximum number of politician parents. Thus, our results do not seem to be driven by influential outliers.

Another potential problem concerns the large size of some of the school classes. Most of our unique school-program-cohort combinations consist of no more than 30 students, which means that they approximate a school class quite well. However, in a few cases, there are hundreds of students in such a combination, in which case they probably capture

several different school classes. Including those combinations introduces some noise to our independent variable, because we cannot separate the treated individuals from the untreated. So what happens if they are excluded?

We test this in two ways. First, Tables A6 and A7 present results when restricting the sample to school classes including 20–35 students. This specification captures what we usually describe as a standard class in upper secondary school. This is an informative specification and we would be worried if we do not find any effect in this specification when we have restricted the sample to a normal class size. Second, Figure A2 shows how our preferred estimates (columns 3 and 6 in the main result tables) are affected if we successively exclude observations based on the size of the school class by decreasing the maximum school class size from 300 to our start out point of 25 students which we use to calculate the variable of interest in our main analysis. The smallest number of students is set to 5 students per class in this specification.[1] It is clear from both Tables A6 and A7 and from Figure A2 that our main findings are not very sensitive to restricting the maximum school class size. As expected, the estimates become noisier when only including individuals attending smaller classes in the estimation in Figure A2. However, the point estimates are consistently positive and do not stray far from the corresponding estimates presented in Tables 2 and 3. It should however be noted that the estimated coefficients are larger in Tables A6 and A7 in comparison to the results in Tables 2 and 3 but less precisely estimated, which is exactly what we would expect given that the sample has been reduced to normal class sizes. Hence, we can conclude that our main findings do not appear to hinge on our choice to also retain the larger school classes.

[1]In other words, this means that we exclude students attending larger schools, most often in larger municipalities, in which the more popular programs are divided into several classes. Also note that the sample becomes very selective for the lower ranges. Some of the included classes in this case have to be special because it is fairly uncommon to have such small classes in upper secondary school.

Table A6: Voter turnout – restricted to classes between 20 and 35 students

	(1) Vote09	(2) Vote09	(3) Vote09	(4) Vote10	(5) Vote10	(6) Vote10
Number of politicians	0.438	0.363	0.368	0.169	0.181	0.213
	(0.136)	(0.141)	(0.141)	(0.121)	(0.121)	(0.117)
Mean dep.var.	33.480	33.935	34.446	76.392	77.040	77.915
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.065	0.076	0.097	0.076	0.081	0.139
Observations	190,363	173,570	170,189	225,744	207,023	203,564

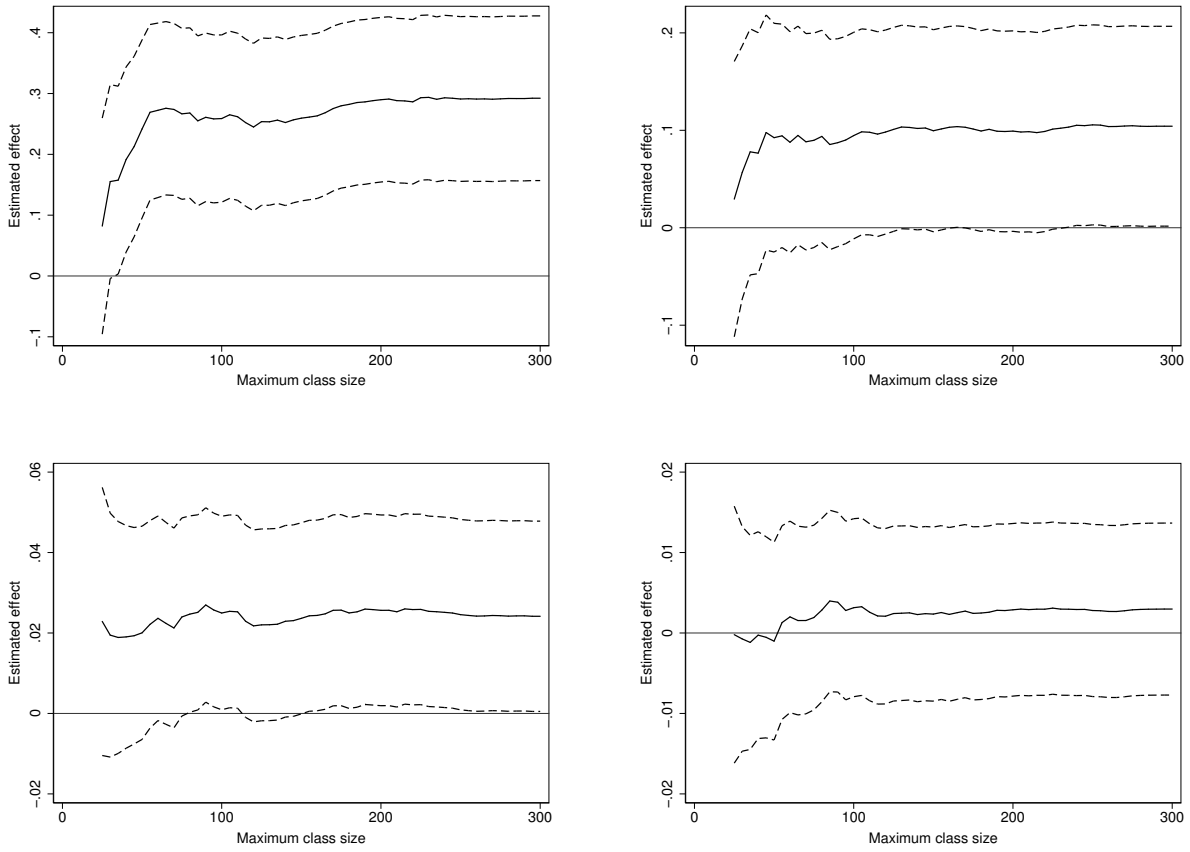
Note: Results from models estimated on a sample restricted to students attending classes of size 20–35 students. The outcome in columns 1–3 is turnout in the 2009 EP election whereas estimates for turnout in the 2010 national election are presented in columns 4–6. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A7: Nominated and elected – restricted to classes between 20 and 35 students

	(1) Nom	(2) Nom	(3) Nom	(4) Elec	(5) Elec	(6) Elec
Number of politicians	0.027	0.025	0.023	−0.002	0.001	0.002
	(0.024)	(0.024)	(0.024)	(0.009)	(0.009)	(0.010)
Mean dep.var.	0.560	0.557	0.558	0.096	0.096	0.097
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.004	0.004	0.006	0.000	0.000	0.001
Observations	228,201	208,581	205,111	228,201	208,581	205,111

Note: Results from models estimated on a sample restricted to students attending classes of size 20–35 students. The outcome in columns 1–3 is running for office at least once in the five elections held between 1998 and 2010 whereas columns 4–6 instead display results for winning office at least once in the same elections. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Figure A2: Estimated effects with sample restrictions on class size



Note: The graphs display coefficient estimates (solid line) and 95% confidence intervals (dashed lines) from models in which the sample is restricted such that school classes surpassing the specified maximum size on the x-axis (between 25 and 300) are excluded. The smallest number of students is always set to 5.

Restricting the treatment to elected politicians

When we calculate our independent variable, we count anyone as a politician who ran for office during the previous election (was nominated). Our reason for including also the ones who were not elected is that a majority of them still serve on different municipal boards and committees. In Tables A8 and A9 we show how the results change if we only include the politicians who were elected.

Overall, the estimates are well in line with the results presented in Tables 2 and 3 in the main text. Reflecting the smaller amount of variation in the treatment variable the coefficients are less precisely estimated. The point estimates for turnout in the 2009 EP and the 2010 general elections (Table A8) are somewhat larger whereas the positive effects of vertical ties on running for office are very similar to the ones obtained when not restricting the the treatment to elected politicians. The point estimates for the most demanding outcome – being elected – are close to zero and very imprecisely estimated.

Table A8: Voter turnout – elected politician parents

	(1) Vote09	(2) Vote09	(3) Vote09	(4) Vote10	(5) Vote10	(6) Vote10
Number of politicians	0.523	0.413	0.394	0.336	0.265	0.201
	(0.122)	(0.124)	(0.123)	(0.098)	(0.099)	(0.094)
Mean dep.var.	39.466	40.088	40.815	79.577	80.271	81.256
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.078	0.093	0.118	0.065	0.070	0.135
Observations	1,020,727	932,262	909,690	1,214,146	1,114,349	1,091,336

Note: Results from models restricting the treatment variable to parents who occupied a political office while their children attended upper secondary school. The outcome in columns 1–3 is turnout in the 2009 EP election whereas estimates for turnout in the 2010 national election are presented in columns 4–6. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A9: Nominated and elected – elected politician parents

	(1) Nom	(2) Nom	(3) Nom	(4) Elec	(5) Elec	(6) Elec
Number of politicians	0.014	0.023	0.026	−0.004	−0.003	−0.002
	(0.020)	(0.021)	(0.021)	(0.009)	(0.009)	(0.009)
Mean dep.var.	0.615	0.610	0.615	0.114	0.113	0.114
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.003	0.003	0.004	0.001	0.001	0.001
Observations	1,226,245	1,122,179	1,099,139	1,226,245	1,122,179	1,099,139

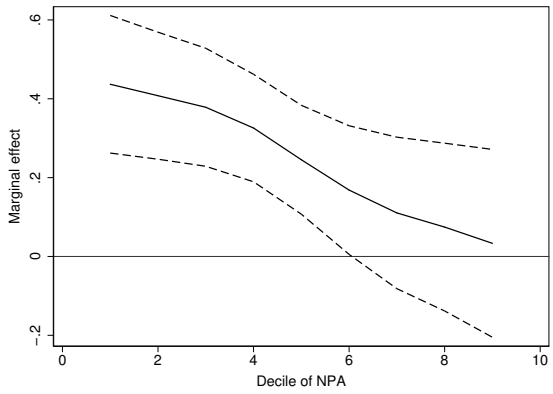
Note: Results from models restricting the treatment variable to parents who occupied a political office while their children attended upper secondary school. The outcome in columns 1–3 is running for office at least once in the five elections held between 1998 and 2010 whereas columns 4–6 instead display results for winning office at least once in the same elections. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Marginal effects over different deciles of NPA

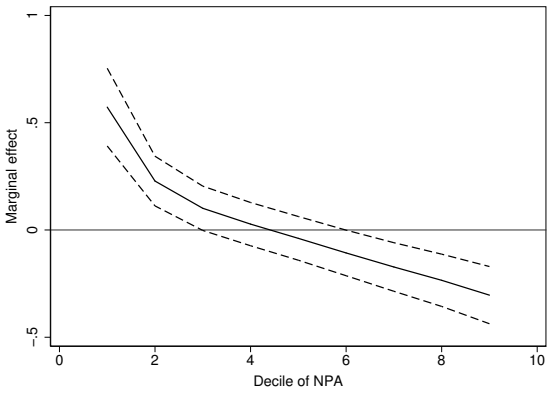
In the main text we present graphs displaying results from a flexible regression model in which a cubic spline function of nascent political ambition (NPA) is interacted with our measure of vertical political ties in Figure 6. In these models we use a spline regression with 5 knots. The rationale behind this modeling strategy is the ability of the spline regression to pick up possible non-linear treatment heterogeneity. However, there are of course other ways to test for treatment heterogeneity. Figure A4 shows results from a linear model where we simply interact our treatment variable with the measure of NPA. The marginal effects are plotted over different deciles, similar to how the splines were presented in the main text.

The overall pattern of the coefficients is well in line with the findings reported in the main text in Figure 6. Clearly, the positive effect of vertical political ties on voting is most marked in the bottom of the NPA distribution. Turning instead to our two measures of elite participation, the opposite holds true. The effect of one more politician per 25 students on running for and winning office is increasing in NPA.

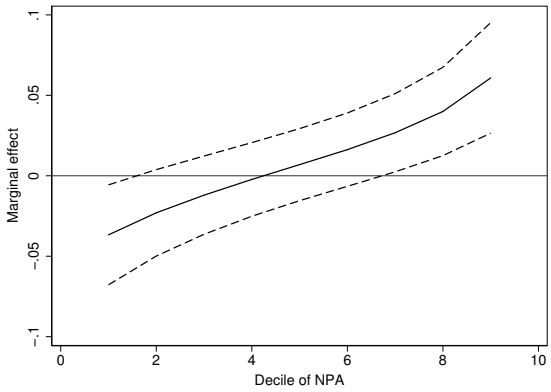
Figure A4: Marginal effect over different deciles of NPA



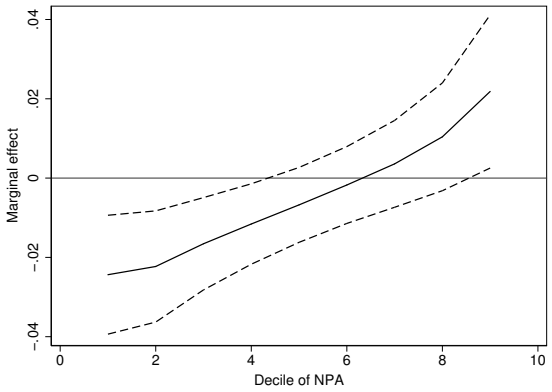
(a) Voting in 2009



(b) Voting in 2010



(c) Ever nominated



(d) Ever elected

Linear interaction with years since graduation

Tables A10 and Table A11 present results from models in which the treatment effect is interacted with a linear measure of the number of years since graduation up until 2009 (voter turnout in the EP election) and 2010 (the other outcomes). In all models the year since graduation variable is recoded such that 1 equals sample maximum (12 years for 2009, 13 years for 2010).

The multiplicative interaction coefficients are positive in all specifications and sometimes statistically significant implying that the treatment effect becomes positive and grow stronger the longer time has passed since graduation. It is worth noting that the interacted treatment effect is substantially larger in the older cohorts and that the positive estimated effect presented in the main analysis is driven by these older cohorts. This would be in line with our findings in the mechanism section in the main text that the treatment effect is mediated by the children of politician parents.

Table A10: Linear interaction with years since graduation: Voter turnout

	(1) Vote09	(2) Vote09	(3) Vote09	(4) Vote10	(5) Vote10	(6) Vote10
Number of politicians	0.115	−0.035	0.140	−0.013	−0.073	−0.011
	(0.137)	(0.138)	(0.136)	(0.106)	(0.107)	(0.098)
Years grad 2009 * N. Pol.	0.401	0.592	0.290			
	(0.221)	(0.221)	(0.216)			
Years grad 2010 * N. Pol.				0.267	0.321	0.231
				(0.175)	(0.174)	(0.157)
Mean dep.var.	39.466	40.088	40.815	79.577	80.271	81.256
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.078	0.093	0.118	0.065	0.070	0.135
Observations	1,020,727	932,262	909,690	1,214,146	1,114,349	1,091,336

Note: Results from OLS regressions. Standard errors, shown in parentheses, allow for clustering at the school-program level. The years since graduation dummy is not included given that we already add cohort fixed effects.

Table A11: Linear interaction with years since graduation: Elite participation

	(1) Nom	(2) Nom	(3) Nom	(4) Elec	(5) Elec	(6) Elec
Number of politicians	−0.043	−0.039	−0.061	−0.015	−0.013	−0.020
	(0.018)	(0.018)	(0.018)	(0.009)	(0.008)	(0.009)
Years grad 2010 * N. Pol.	0.127	0.125	0.168	0.030	0.032	0.045
	(0.034)	(0.035)	(0.036)	(0.016)	(0.016)	(0.017)
Mean dep.var.	0.615	0.610	0.615	0.114	0.113	0.114
Individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Parent covariates	No	Yes	Yes	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	Yes	No	No	Yes
Adjusted R2	0.003	0.003	0.004	0.001	0.001	0.001
Observations	1,226,245	1,122,179	1,099,139	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. Age refers to average age across the elections the individual was eligible to run in. Standard errors, shown in parentheses, allow for clustering at the school-program level. The years since graduation dummy is not included given that we already add cohort fixed effects.

Additional mechanism analyses

In the mechanism section we also discuss a set of analyses intended to test if the treatment effect is mediated by intermediary mechanisms, such as changes in students' psychological engagement and recruitment activities.

First we present a table testing whether having a child in a class where there are politicians parents affects the probability that the non-politician parent is more likely to run for office in the future after the child finished upper secondary school. The results are presented in Table A12. We have run an analysis with our standard treatment variable and divided the treatment variable into the number of politicians parents in a class among the students that are interested in politics and those who are not interested in politics (proxied by having voted in the 2009 European Parliament election). The outcome is defined as being nominated after the child has attended upper secondary school and is separated between fathers and mothers. All of the estimates in Table A12 are small and statistically significant indicating that the non-politician parents are not affected. We may as a result rule out that the estimated effect in the main text captures an intergenerational transmission taking place after upper secondary school.

In Tables A13 through A19 we display estimates from a number of separate party-specific models in which we employ two treatments – the number of politician parents running for a specific party (e.g. the Social Democrats) and the number of politicians running for other parties – and the outcome is running for the same party among the children. Most of the estimates in these tables are small in magnitude and statistically insignificant.

Finally, Table A20 reports estimated effects of our treatment variable on individuals' self-reported willingness to take part in political discussions as measured by the yearly Living Conditions Surveys (ULF/SILC) between 1997 and 2010.[2] The sample size in these analyses is much smaller than the ones used in the main analyses and the estimates are, consequently, less precise. Nevertheless, we can see that in all specifications vertical social ties to politicians during upper secondary school, although not statistically significant, have a positive effect on individuals' willingness to engage in political discussions.

[2]The variable is missing for 2006. We use the first year if an indiviudal has appear several times in the panel.

Table A12: Does the effect goes through the non-politician parents?

	(1) Father.Nom	(2) Mother.Nom	(3) Father.Nom	(4) Mother.Nom
Number of politicians	0.000	0.000		
	(0.000)	(0.000)		
Number of politicians vot. stud.			0.000	−0.000
			(0.000)	(0.000)
Number of politicians n.v stud.			−0.000	0.000
			(0.000)	(0.000)
Mean dep.var.	0.007	0.006	0.007	0.006
Individual covariates	No	No	No	No
Parent covariates	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Muni. FE	Yes	Yes	Yes	Yes
Adjusted R2	0.005	0.005	0.005	0.005
Observations	1,099,139	1,099,139	1,099,139	1,099,139

Note: Results from OLS regressions. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A13: Partisan mobilization: Moderate Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	0.027	0.037	0.034
	(0.030)	(0.032)	(0.032)
Number of politcians other	0.000	−0.003	−0.006
	(0.009)	(0.010)	(0.010)
Mean dep.var.	0.447	0.467	0.465
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.002	0.003	0.004
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Moderate Party (MP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A14: Partisan mobilization: Christian Democratic Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	0.025	0.019	0.016
	(0.027)	(0.028)	(0.029)
Number of politcians other	0.007	0.008	0.009
	(0.009)	(0.009)	(0.009)
Mean dep.var.	0.251	0.266	0.266
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.003	0.004	0.005
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Christian Democratic Party (CDP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A15: Partisan mobilization: Liberal Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	0.016	0.012	0.016
	(0.028)	(0.030)	(0.030)
Number of politcians other	0.003	0.003	0.002
	(0.006)	(0.007)	(0.007)
Mean dep.var.	0.216	0.226	0.225
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.001	0.002	0.002
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Liberal Party (LibP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A16: Partisan mobilization: Center Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	0.032	0.033	0.039
	(0.028)	(0.030)	(0.031)
Number of politcians other	0.020	0.023	0.023
	(0.012)	(0.013)	(0.013)
Mean dep.var.	0.458	0.479	0.482
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.005	0.006	0.007
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Center Party (CP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A17: Partisan mobilization: Green Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	−0.064	−0.072	−0.074
	(0.021)	(0.022)	(0.023)
Number of politcians other	0.009	0.008	0.009
	(0.006)	(0.006)	(0.006)
Mean dep.var.	0.143	0.149	0.148
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.001	0.002	0.002
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Green Party (GP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A18: Partisan mobilization: Social Democrats

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	0.003	0.005	0.005
	(0.024)	(0.025)	(0.025)
Number of politcians other	0.007	0.009	0.007
	(0.015)	(0.016)	(0.016)
Mean dep.var.	0.691	0.725	0.723
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.003	0.003	0.004
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Social Democrats (SP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A19: Partisan mobilization: Left Party

	(1) Nom	(2) Nom	(3) Nom
Numer of politicians SP	−0.014	−0.013	−0.014
	(0.025)	(0.027)	(0.027)
Number of politcians other	−0.002	−0.002	−0.001
	(0.007)	(0.008)	(0.008)
Mean dep.var.	0.222	0.231	0.231
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.001	0.002	0.003
Observations	1,226,245	1,122,179	1,099,139

Note: Results from OLS regressions. The outcome in columns 1–3 is running for office for the Left Party (LP) at least once in the five elections held between 1998 and 2014. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A20: Treatment effects on political discussion

	(1) PolDisc	(2) PolDisc	(3) PolDisc
Number of politicians	0.011	0.009	0.013
	(0.012)	(0.013)	(0.013)
Mean dep.var.	0.426	0.427	0.425
Individual covariates	Yes	Yes	Yes
Parent covariates	No	Yes	Yes
Cohort FE	Yes	Yes	Yes
Muni. FE	No	No	Yes
Adjusted R2	0.034	0.041	0.045
Observations	8,372	7,698	7,582

Note: Results from OLS regressions. The outcome in columns 1–3 is discussing politics. Standard errors, shown in parentheses, allow for clustering at the school-program level.

Main results tables: 5 columns with additional specifications

In Tables 2–3 in the main text we present three columns for each outcome variable (six columns in total in each table) for space reasons. Here we add additional specifications for transparency by presenting four tables with five columns in each. We have separated the vector for parental covariates into two parts where we add parental covariates for individual i separately from parental covariates for the class c.

Table A21: European election voter turnout

	(1)	(2)	(3)	(4)	(5)
Number of politicians	0.346	0.328	0.327	0.279	0.294
	(0.067)	(0.067)	(0.069)	(0.069)	(0.069)
Mean dep.var.	39.465	39.466	40.088	40.088	40.815
Individual covariates	No	Yes	Yes	Yes	Yes
Parent covariates i	No	No	Yes	Yes	Yes
Parent Class covariates	No	No	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	No	No	Yes
Adjusted R2	0.072	0.078	0.093	0.093	0.118
Observations	1,020,768	1,020,727	932,262	932,262	909,690

Note: Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A22: National election voter turnout

	(1)	(2)	(3)	(4)	(5)
Number of politicians	0.141	0.122	0.119	0.089	0.105
	(0.056)	(0.055)	(0.056)	(0.056)	(0.052)
Mean dep.var.	79.574	79.577	80.271	80.271	81.256
Individual covariates	No	Yes	Yes	Yes	Yes
Parent covariates i	No	No	Yes	Yes	Yes
Parent Class covariates	No	No	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	No	No	Yes
Adjusted R2	0.055	0.065	0.070	0.070	0.135
Observations	1,214,188	1,214,146	1,114,349	1,114,349	1,091,336

Note: Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A23: Nominated

	(1)	(2)	(3)	(4)	(5)
Number of politicians	0.022	0.022	0.024	0.024	0.024
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Mean dep.var.	0.615	0.615	0.610	0.610	0.615
Individual covariates	No	Yes	Yes	Yes	Yes
Parent covariates i	No	No	Yes	Yes	Yes
Parent Class covariates	No	No	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	No	No	Yes
Adjusted R2	0.003	0.003	0.003	0.003	0.004
Observations	1,226,287	1,226,245	1,122,179	1,122,179	1,099,139

Note: Standard errors, shown in parentheses, allow for clustering at the school-program level.

Table A24: Elected

	(1)	(2)	(3)	(4)	(5)
Number of politicians	0.001	0.001	0.002	0.003	0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Mean dep.var.	0.114	0.114	0.113	0.113	0.114
Individual covariates	No	Yes	Yes	Yes	Yes
Parent covariates i	No	No	Yes	Yes	Yes
Parent Class covariates	No	No	No	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Muni. FE	No	No	No	No	Yes
Adjusted R2	0.001	0.001	0.001	0.001	0.001
Observations	1,226,287	1,226,245	1,122,179	1,122,179	1,099,139

Note: Standard errors, shown in parentheses, allow for clustering at the school-program level.

References

Lindgren, Karl-Oskar, Sven Oskarsson, and Mikael Persson (2017). *Can increased education help reduce the political opportunity gap?* Working Paper 2017:12. Institute for Evaluation of Labour Market and Education Policy (IFAU).

Stammann, Amrei, Florian Heiß, and Daniel McFadden (2016). *Estimating fixed effects logit models with large panel data*. Conference Paper. Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016.