Online Appendix for "Educating for Democracy? Going to College Increases Political Participation"

Contents

Appendix A: Additional Tables

A1. Interaction Analyses Testing the Compensation Hypothesis

Table A1 provides the regression results that underlie the test of the compensation hypothesis reported in the results section of the article. Moreover, I test the so-called compensation hypothesis which posits that the effect of education is larger among less privileged students as it compensates for lack of political socialization at home (Neundorf, Niemi, and Smets 2016; Lindgren et al. 2019; see also Appendix J). In a recent study, Mendelberg et al. (2021) found that having more affluent peers in college increased political participation but not in a compensatory way. The specification of these regressions is identical to the main two-way fixed-effects model reported in the article, except I add to equation (1) an interaction between the time period indicator, the treatment variable, and the subgroup-defining variable. In table A1, the subgroup-defining variable is whether the respondent voted in the pre-treatment election in 2004 or not

TABLE A1 *Compensation for Previous Participation: Heterogenous Effects of College Education on Voter Turnout by Prior Participation*

	Full Sample (interaction)		
	(1)	(2)	(3)
No College	ref.	ref.	ref.
Attended College	0.097***	0.087***	0.059*
	(0.022)	(0.024)	(0.025)
Attended College × 2004-voter	ref.	ref.	ref.
Attended College × 2004 non-voter	0.065*	0.070*	0.078*
	(0.029)	(0.031)	(0.031)
Time FEs & Individual FEs	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			✓
Units	10,426	9,652	9,554
Observations	19,813	18,298	18,118

Note: Interaction model testing the difference in effect size between those who had previously turned out to vote and those who had not. If the estimated interaction is positive, it means that the effect is larger among previous non-voters. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. All models include time trends that are specific to whether a respondent voted in the pre-college election or not. Model 3 further interacts time fixed effects with pre-college cognitive skills, gender, race and parental education and income. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

(2004 voter). Note that the statistical power in the interaction analysis is somewhat reduced due to the added interaction terms and the fact that the untreated group is already small compared to the treated group. However, we can confidently conclude that the estimated effect is larger among previous non-voters compared to previous voters by observing the positive and statistically significant interaction terms in Table A1 ($p \leq 0.025$). Moreover, tables A1A and A1B further test the compensation hypothesis, but uses self-reported race and parental socio-economic status as the subgroup defining variable. Table A1A indicates that the effect of college was in fact larger for Black and Hispanic students than for White or Asian students. In table A1B, while the estimated interaction term for being in the bottom quartile of parental SES is indeed positive, it is substantively smaller and does not reach statistical significance in the most restrictive model. The evidence of socio-economic compensation is therefore mixed.

Whereas the findings regarding prior voting does not replicate in the NELS:88 replication sample (cf. Appendix B), the findings for race and socio-economic status are quite similar across samples. The fact that we only find consistent evidence of a compensation effect in the main sample is puzzling. One explanation may be, as hypothesized in the article, that the relationship between college and voter turnout is changing over time – including the compensating role of higher education. Another may be that we have left to discover what it is that higher education exactly compensates for. In both cases, this puzzling result does motivate further studies of the compensating role of college and heterogeneous effects more broadly.

TABLE A1A *Heterogeneous Effects of College Education on Voter Turnout by Self-Reported Race*

	Full Sample (interaction)		
	(1)	(2)	(3)
Attended College × White/Asian	ref.	ref.	ref.
Attended College × Black/Hispanic/Other	0.085***	0.091***	0.079***
	(0.013)	(0.014)	(0.014)
Time FEs & Individual FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			✓
Units	11,131	9,667	9,569
Observations	20,518	18,313	18,133

Note: Interaction model testing the difference in effect size between groups defined by self-reported race. If the estimated interaction is positive, it means that the effect is larger among black and hispanic voters. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. Model 3 interacts time fixed effects with pre-college cognitive skills, gender and parental education and income. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

TABLE A1B *Heterogeneous Effects of College Education on Voter Turnout by Socio-economic Background*

	Full Sample (interaction)		
	(1)	(2)	(3)
Attended College × High Parental SES (Q2 or above)	ref.	ref.	ref.
Attended College × Bottom Quartile Parental SES	0.056***	0.056***	0.027
	(0.015)	(0.016)	(0.017)
Time FEs & Individual FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			✓
Units	11,127	9,667	9,569
Observations	20,511	18,313	18,133

Note: Interaction model testing the difference in effect size between groups defined by parents' socio-economic status. If the estimated interaction is positive, it means that the effect is larger among voters whose parents are in the bottom quartile of socio-economic status. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. Socio-economic status is an index of parental education and income. Model 3 interacts time fixed effects with pre-college cognitive skills, gender and race. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

A2. Placebo Analysis Probing the Parallel Trends Assumption

Table A2 provides the regression results from the placebo analysis designed to indirectly test the parallel trends assumption. This analysis is presented in the section "Plausibility of Parallel Trends" of the article. Specifically, to indirectly test the parallel trends assumption, I use variation in future college attendance within the group who does not attend college between 2004 and 2008 ("untreated group"). The placebo difference-in-differences reported in the table thus test whether the pre-college trends in turnout were different for future college-attendees compared to future non-attendees. I find an insignificant difference in pre-college trends for these groups across specifications, with our most strict specifications yielding point estimates of 1.9 percentage points ($p = 0.59$) and 2.7 percentage points ($p = 0.45$). This indicates that turnout trends from 2004 to 2008 were not different for future college-attendees and future non-attendees.[1] However, as the main article notes, there are important limitations to this kind of placebo-analysis, which means that while we fail to reject that turnout trends were parallel for future college attendees and future non-attendees in this test, we cannot confidently confirm it either (Rambachan and Roth 2023). We may still regard the placebo difference-in-differences as a point estimate of the amount of selection bias in the main effect estimates. Following a logic similar to that of Rambachan & Roth (2023), we might assume that there is actually a violation of the parallel trends assumption equal to this number. Under this assumption, however, we can still confidently rule out that this selection bias makes up the entire estimated effect. This follows from observing that the assumed violation of 1.9 pp., 2.7 pp., and 5.2 pp. for the different specifications is well below the lower bound of the 95% confidence intervals of the effect estimates in all cases except model 4 of Table 1 in the main article. Specifically, the confidence intervals are $[10.9; 16.5]$, $[9.9; 16.1]$, $[7.5; 13.9]$, $[4.1; 15.1]$, $[4.5; 16.7]$, and $[2.8; 15.0]$ in models 1, 2, 3, 4, 5, and 6, respectively (cf. Table 1 in the main article). More specifically, the true violation of the parallel trends assumption would have to be larger in magnitude than the placebo estimate by at least a factor of 1.6 in order for the most conservative

---

[1] It should be noted that if we exclude both time-varying covariates and covariate-by-time fixed effects (model 3 of Table A2), we arrive at a point estimate of 5.3 percentage points ($p = 0.10$). However, in the placebo analysis, we do not really have good reasons to exclude the time-varying covariates like in the main analyses: They are not problematic here since they are not post-treatment variables. This means that part of this larger placebo estimate, compared to models 1-2, likely is driven by the life events that we control for in models 1 and 2 but not in model 3. This does, however, suggest that we should put less weight on model 1 in Table 1 of the main article as we are not as confident in the parallel trends assumption for this specification, based on the placebo analyses. Importantly, the results are very similar and consistent across models – also excluding that model.

and statistically uncertain of our effect estimates to be statistically insignificant, i.e. model 6 of Table 1 (Rambachan and Roth 2023). Therefore, even if there were violations as large as 1.6 times the placebo-estimates, there would still be a positive effect. This arguably strengthens the causal credibility of the findings.

Besides providing an indirect test of the parallel trends assumption in general, the placebo analyses also have implications – as noted in the paper – for which of the analyses in Table 1 of the main article should be trusted the most, in relative terms. There are two important observations in this regard: one regarding which control specification is preferable, and one regarding which comparison group is preferable. Firstly, regarding the control specification, we see pretty clearly that the more extensively controlled specifications (e.g. including pre-college covariates-by-time fixed effects) yields the smallest placebo estimate. This means that when comparing specifications in Table 1 *that use the same comparison group,* we should prefer the models with extensive controls – i.e. we should prefer models 2 and 3 to model 1 and similarly prefer models 6 and 5 to model 4.

Secondly, the implications for which comparison group is preferable are just as important. Note that the placebo estimates roughly equal the decrease in effect size between the full sample models (models 1–3) and the future college-attendees models (models 4-6) in Table 1 of the article. This suggests that this latter, more restrictive comparison group should be preferred. In this regard, it is relevant to highlight that the placebo models compare future college-attendees to never attendees. This means that the placebo analyses speak most directly to the full sample models (models 1–3 in Table 1 of the article). This is because models 1–3 in Table 1 of the main article compare college-goers to both future college-attendees *and* never-attendees, whereas the models 4–6 only compare college-goers to future college-attendees. In conjunction with the reduction in effect estimates when restricting the comparison to future college-attendees in the main analyses, this might suggest that the placebo estimates would likely be even smaller in absolute terms if we were able to perform placebo analyses for this comparison group.

Finally, it should be noted that the placebo analysis presented here is only based on a subsample of respondents. This is because the cohort of respondents in this study were simply not eligible to vote in elections before 2004 due to their age, and it is, therefore, not possible to make a typical pre-trend analysis for the entire sample using pre-treatment data. At the election in 2000, the respondents were only 14 years old.

TABLE A2 *Placebo-test of parallel trends within the "No college" group of the main analyses: Comparing 2004-2008 turnout trends of future college students to those who never went to college*

	(1)	(2)	(3)
No College	ref.	ref.	ref.
Attended College	0.019	0.027	0.053
	(0.035)	(0.035)	(0.032)
Time FEs & Individual FEs	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓
Time-varying Controls	✓	✓	
Additional Pre-college Covariates × Time FEs	✓		
Units	1,176	1,206	1,385
Observations	2,166	2,216	2,572

Note: Placebo difference-in-differences estimates of how future college-goers changed their turnout between the 2004 and 2008 elections, compared to respondents who never went to college. Robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. All models include time trends that are specific to whether a respondent voted in the 2004 election or not.

A3. Full Results for Main Article's Table 1 (Estimates for all Covariates)

TABLE A3 *Full Results: Effect of College Education on Voter Turnout*

	Full Sample			Restricted Ctrl.grp			Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main Difference-in-differences Estimates:								
No College × PostPeriod = 1	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Attended College × PostPeriod = 1	0.137 (0.014)	0.130 (0.016)	0.107 (0.016)	0.096 (0.028)	0.106 (0.031)	0.089 (0.031)	0.116 (0.026)	0.101 (0.027)
Covariates:								
PostPeriod = 0 (2004)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
PostPeriod = 1 (2008)	0.308 (0.013)	0.278 (0.018)	0.291 (0.024)	0.355 (0.028)	0.301 (0.033)	0.308 (0.037)	-0.040 (0.026)	0.313 (0.030)
Moved from 2004 Residence		0.024 (0.011)	0.026 (0.011)		0.029 (0.012)	0.029 (0.012)		0.022 (0.013)
Married		-0.054 (0.034)	-0.042 (0.034)		-0.070 (0.038)	-0.057 (0.038)		-0.064 (0.038)
Living with Parents		0.033 (0.010)	0.011 (0.010)		0.038 (0.010)	0.013 (0.011)		0.028 (0.011)
Parental Divorce		0.007 (0.019)	-0.002 (0.019)		0.007 (0.020)	0.002 (0.020)		0.005 (0.024)
Parental Job Loss		-0.006 (0.014)	-0.002 (0.014)		0.001 (0.014)	0.004 (0.014)		-0.016 (0.017)
Parent Died		-0.056 (0.035)	-0.045 (0.034)		-0.091 (0.037)	-0.078 (0.036)		-0.046 (0.041)
Friend Died		0.012 (0.010)	0.009 (0.010)		0.014 (0.010)	0.011 (0.010)		0.010 (0.011)
Became Seriously Ill or Disabled		0.019 (0.024)	0.014 (0.024)		0.020 (0.025)	0.015 (0.024)		0.025 (0.028)
Serious Ilness in The Family		0.004 (0.012)	-0.001 (0.011)		0.003 (0.012)	-0.001 (0.012)		0.008 (0.013)
Violent Crime Victimization		0.055 (0.032)	0.053 (0.031)		0.041 (0.033)	0.042 (0.032)		0.039 (0.035)
Became Parent		-0.005 (0.029)	-0.014 (0.028)		0.009 (0.033)	0.000 (0.032)		-0.016 (0.042)
Employed		-0.003 (0.010)	-0.004 (0.010)		-0.006 (0.010)	-0.007 (0.010)		-0.008 (0.011)

(Table A3 is continued on the next page…)

(Table A3, continued from previous page)	Full Sample			Restricted Ctrl.grp			Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-college Covariates × Time FEs								
Pre-college Non-Voter × PostPeriod = 1	ref.	ref.	ref.	ref.	ref.	ref.		ref.
Pre-college Voter × PostPeriod = 1	-0.607 (0.010)	-0.613 (0.010)	-0.632 (0.010)	-0.615 (0.010)	-0.621 (0.011)	-0.642 (0.011)		-0.609 (0.011)
Parental SES 1st quartile × PostPeriod = 1			ref.			ref.		
Parental SES 2nd quartile × PostPeriod = 1			0.023 (0.015)			0.017 (0.016)		
Parental SES 3rd quartile × PostPeriod = 1			0.041 (0.015)			0.033 (0.016)		
Parental SES 4th quartile × PostPeriod = 1			0.090 (0.015)			0.085 (0.016)		
Cognitive Skills 1st quartile × PostPeriod = 1			ref.			ref.		
Cognitive Skills 2nd quartile × PostPeriod = 1			-0.013 (0.016)			-0.009 (0.018)		
Cognitive Skills 3rd quartile × PostPeriod = 1			0.012 (0.016)			0.020 (0.017)		
Cognitive Skills 4th quartile × PostPeriod = 1			0.022 (0.016)			0.030 (0.018)		
Race: Hispanic × PostPeriod = 1			ref.			ref.		
Race: Other × PostPeriod = 1			0.099 (0.018)			0.096 (0.019)		
Race: Asian × PostPeriod = 1			-0.068 (0.021)			-0.069 (0.022)		
Race: White, Non-Hispanic × PostPeriod = 1			-0.002 (0.015)			0.001 (0.016)		
Sex: Female × PostPeriod = 1			ref.			ref.		
Sex: Male × PostPeriod = 1			-0.063 (0.009)			-0.063 (0.010)		
Constant	0.570 (0.002)	0.582 (0.007)	0.585 (0.008)	0.597 (0.002)	0.609 (0.008)	0.611 (0.008)	0.608 (0.003)	0.621 (0.009)
Time FEs & Individual FEs	✓	✓	✓	✓	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Controls		✓	✓		✓	✓		✓
Additional Pre-college Covariates × Time FEs			✓			✓		
Units	10,426	9,652	9,554	9,360	8,717	8,644	9,289	8,582
Observations	19,813	18,298	18,118	17,842	16,593	16,454	18,578	17,068

Note: All estimates from the estimations on which table 1 of the main article is based. The reference categories for the time-varying covariates from "Moved (…)" to "Employed" are simply obtained by adding "not" to the label that the covariate has. E.g. "Not Employed". Models 7 and 8 compare college attendees to a matched comparison group (Genetic Matching with replacement based on pre-college race, gender, prior voting and parental SES). Robust standard errors clustered by individual in parentheses. Significance stars omitted for reasons of space. Note that the dependent variable of all regressions is turnout Y_{it} which covers each respondent (i) at two points in time (t), namely both in 2004 and 2008.

It is important to note that due to the individual-level fixed effects in all models, any coefficient that may be obtained by adding time-invariant covariates to the regressions would be absorbed by the individual fixed effects. This is true for, e.g., a lower-order term for the “Attended College” variable. The time-invariant differences in turnout between those who attend college and those who do not are simply already accounted for and subsumed in the individual-level fixed effects. The same applies to the lower-order term for the “Pre-college Voter” variable. Also note that the dependent variable of all regressions is turnout Y_{it} which covers each respondent (i) at two points in time (t), namely both in 2004 and 2008.

It is worth stating the interpretation of some of the key coefficients in the simplest model, namely model (1): The coefficient for the PostPeriod-indicator indicates that among non-attendees who did not vote in 2004, there was an average increase in turnout from 2004 to 2008 of about 31 percentage points (recall that this group of pre-college non-voters is defined by starting from 0 per cent in 2004). The coefficient for the interaction between Pre-college Voting and PostPeriod-variables (appearing first on the second page of the table) indicates that this change over time is different by about 61 percentage points when compared to non-attendees who *did* vote in 2004. This means that among non-attendees who did vote in 2004, there was an average *decrease* in turnout from 2004 and 2008 of about 30 ($= 31\% - 61\%$) percentage points (recall that this group of pre-college voters is defined by starting from 100 per cent in 2004). Finally, the coefficient of interest (Attended College × PostPeriod) indicates how these two trends are different when altering only whether people attended college. Looking at pre-college non-voters, turnout increased by an additional 13.7 percentage points among college attendees compared to non-attendees. This means that, among pre-college non-voters, non-attendees experienced an increase in turnout from 2004 to 2008 of about 31 percentage points, while college attendees experienced an increase in turnout from 2004 to 2008 of about 45 percentage points (recall again they were all characterized by starting from 0 per cent). Conversely, among pre-college voters: While non-attendees in this group experienced a decrease in turnout of about 30 percentage points, college attendees in this group only experienced a drop of about 16 percentage points ($= -30\% + 13.7\%$).

Appendix B: Replicating the Analyses in an Independent Sample

I use data from an independent panel sample from a cohort 12 years prior to the ELS:2002 to replicate the main analyses, namely the National Education Longitudinal Study of 1988 (NELS:88, see Appendix B2). Using this dataset, I conduct both an estimation of the average effect of college attendance on turnout and a test of the compensation hypothesis. This independent sample yields estimates that corroborate the finding of an overall effect of college in the main sample. However, we do not find consistent evidence of a compensation effect in this sample as we do in the main sample. I detail both sets of results below.

Table B1 shows the difference-in-differences analyses for the average effect of college attendance on turnout. We estimate that attending college leads to an increase in voter turnout of between 12.6 and 13.7 percentage points in this sample (all $p < 0.001$). These estimates are remarkably similar to those obtained in the full sample models (models 1-3) in the ELS:2002 sample, reported Table 1 in the article. For reference, these estimates are 12.8 and 13.7 percentage points. When restricting the untreated group to future college-attendees (models 4-6), the point estimates are slightly reduced – just as it is the case in the ELS:2002 models. Finally, the matched comparison group models (models 7 and 8) yield estimates ranging from 13.2 and 14.9 percentage points. In the restricted models (4-6) and matched models (7-8) of Table B1, the effect estimates are about 3-4 percentage points larger than in the corresponding ELS:2002 sample models. This larger effect size may be due to the higher retention rate of the NELS:88 sample. This squares well with the idea mentioned in the main paper (and appendix H), that non-response may attenuate the effect of education on turnout (since non-responders are likely to be both low-education and low-turnout individuals). Therefore, the ELS:2002 study may be more likely to produce a conservative estimate.

Table B2 provides the regression results from the interaction analyses testing the compensation hypothesis in the replication sample. The specification of these regressions is, again, identical to the main two-way fixed-effects model reported in the article, except I add to equation (1) an interaction between the time period indicator, the treatment variable, and the subgroup-defining variable. In Table B2 the subgroup defining variable is whether the respondent voted in the pre-treatment election or not. In the replication sample, the pre-treatment election year is 1992 (post-treatment is 1996). We do not find evidence of a difference in effect sizes between previous voters and previous non-voters. The three models estimate a difference in effect sizes between 0.7 and 1.6 percentage points ($p \geq 0.49$). Tables B2A and B2B further test the compensation hypothesis but uses self-reported race and parental socio-economic status as the subgroup defining variable. Table B2A indicates that the effect of college was in fact larger for Black and Hispanic students than for White or Asian students. The evidence of socio-economic compensation in Table B2B is more mixed, yielding a substantively smaller interaction term that does not reach statistical significance in the most restrictive model.

The fact that we only find consistent evidence of a compensation effect in the main sample is puzzling. One explanation may be, as hypothesized in the article, that the relationship between college and voter turnout is changing over time – including the compensating role of higher education. Another may be that we have left to discover what it is that higher education exactly compensates for. In both cases, this puzzling result does motivate further studies of the compensating role of college and heterogeneous effects more broadly.

TABLE B1 *NELS:88 Replication Sample - Effect of College Education on Voter Turnout*

	Full Sample			Restricted Untr.Grp.			Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No College	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Attended College	0.136***	0.137***	0.117***	0.126***	0.134***	0.130***	0.132***	0.149***
	(0.010)	(0.010)	(0.012)	(0.022)	(0.022)	(0.024)	(0.025)	(0.023)
Time FEs & Individual FEs	✓	✓	✓	✓	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Controls		✓	✓		✓	✓		✓
Additional Pre-college Covariates × Time FEs			✓			✓		
Units	10,580	10,414	9,469	8,448	8,328	7,679	9,821	9,691
Observations	21,160	20,828	18,938	16,896	16,656	15,358	19,642	19,382

Note: Difference-in-differences estimates of how college-goers changed their turnout between the 1992 and 1996 elections compared to non-attendees.In models 4-6, the untreated group is restricted to future college-goers. Time-varying controls include residential mobility. All models include time trends that are specific to whether a respondent voted in the pre-college election or not. Models 3 and 6 further interact time fixed effects with pre-college cognitive skills, gender, race and parental education and income. Models 7 and 8 compare college attendees to a matched comparison group (using genetic matching with replacement based on pre-college race, gender, prior voting and parental education and income). OLS estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

TABLE B2 *NELS:88 Replication Sample - Compensation for Previous Participation: Heterogeneous Effects of College Education on Voter Turnout by Prior Participation*

	Full Sample (interaction)		
	(1)	(2)	(3)
No College	ref.	ref.	ref.
Attended College	0.132***	0.133***	0.107***
	(0.017)	(0.017)	(0.019)
Attended College × 1992-voter	ref.	ref.	ref.
Attended College × 1992 non-voter	0.007	0.006	0.016
	(0.022)	(0.022)	(0.023)
Time FEs & Individual FEs	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			
Units	10,580	10,414	9,469
Observations	21,160	20,828	18,938

Note: Interaction model testing the difference in effect size between those who had previously turned out to vote and those who had not. 1992 is the pre-treatment election year. If the estimated interaction is positive, it means that the effect is larger among previous non-voters. Time-varying controls include residential mobility. All models include time trends that are specific to whether a respondent voted in the pre-college election or not. Model 3 further interacts time fixed effects with pre-college cognitive skills, gender, race and parental SES. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Table B2A *NELS:88 Replication Sample - Heterogeneous Effects of College Education on Voter Turnout by Self-Reported Race*

	Full Sample (interaction)		
	(1)	(2)	(3)
Attended College × White/Asian	ref.	ref.	ref.
Attended College × Black/Hispanic/Other	0.097***	0.097***	0.079***
	(0.014)	(0.014)	(0.016)
Time FEs & Individual FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			✓
Units	10,548	10,383	9,469
Observations	21,096	20,766	18,938

Note: Interaction model testing the difference in effect size between groups defined by race. If the estimated interaction is positive, it means that the effect is larger among black and hispanic voters. Time-varying controls include residential mobility. Model 3 interacts time fixed effects with pre-college cognitive skills, gender and parental SES. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Table B2B *NELS:88 Replication Sample - Heterogeneous Effects of College Education on Voter Turnout by Parental Socio-economic Status*

	Full Sample (interaction)		
	(1)	(2)	(3)
Attended College × High Parental SES (Q2-Q4)	ref.	ref.	ref.
Attended College × Bottom Quartile SES	0.045**	0.048**	0.025
	(0.015)	(0.015)	(0.017)
Time FEs & Individual FEs	✓	✓	✓
Time-varying Controls		✓	✓
Additional Pre-college Covariates × Time FEs			✓
Units	10,469	10,304	9,469
Observations	20,938	20,608	18,938

Note: Interaction model testing the difference in effect size between groups defined by parental socio-economic status. If the estimated interaction is positive, it means that the effect is larger among those with parental SES in the bottom quartile. Time-varying controls include residential mobility. Model 3 interacts time fixed effects with pre-college cognitive skills, gender and parental SES. OLS-estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

B2. About the NELS:88 Sample

Just like the ELS:2002, the National Education Longitudinal Study of 1988 (NELS:88) was collected by the National Center of Education Statistics of the U.S. Department of Education as part of a program of several longitudinal cohort studies (Curtin et al. 2002; NCES 2002). It also comprises a survey component and an administrative transcript component. The survey component consists of five waves collected in 1988, 1990, 1992, 1994, and 2002, respectively. The 1988 wave sampled a nationally representative sample of the cohort of 8^{th} graders in public and private schools. This initial wave was sampled based on a clustered, stratified national probability sample with students sampled within schools. In the second wave (1990), the sample was freshened to be representative of the cohort of high school sophomores that year. In the third wave (1992), the sample was freshened to be representative of the cohort of high school seniors that year. The replication in the current study is based on respondents from this senior cohort who also answered the fourth and fifth waves of the survey (1994 and 2000). Out of the 15,649 eligible respondents in the fifth wave, 12,144 cases were completed (77.6%). The third and fourth waves were conducted two and eight years after scheduled high school graduation, respectively. Finally, the administrative transcript component was collected in 2000 and consists of postsecondary education transcripts.

The key independent variable is college enrollment. Out of the 10,580 cohort members who completed all waves of the survey, 2535 (24.0%) did not go to college between the 1992 and 1998 elections, while 8,045 (76.0%) attended college in the period. Table B3 shows the mean values of the outcome variable (voting) for the two elections across the groups defined by the main independent variable.

TABLE B3 *NELS:88 Sample means on voting in 1992 and 1996 by treatment status*

	No College (mean)	Attended College (mean)	All Respondents (mean)
Voted in 1992	0.30	0.57	0.51
Voted in 1996	0.37	0.64	0.57
	n = 2535	n = 8045	n = 10580

Note: Data from the balanced panel sample of respondents who replied to all waves of the survey.

Appendix C: Details on Existing Studies on College and Participation

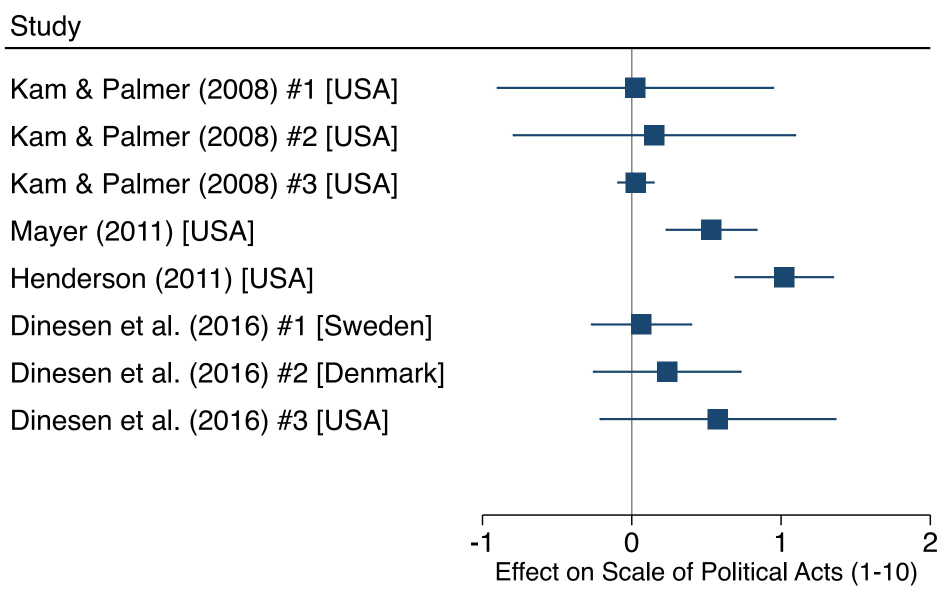
C1. Additional Relevant Literature

C1.A. Studies with Other Participatory Acts as Dependent Variable

As described in the article, previous studies have investigated the effect of college education on political participation. These can be divided into a group of studies that have voter turnout as the dependent variable and a group of studies that use a survey scale of other participatory acts as the dependent variable. Figure 1 of the main article summarizes the results from the former group estimating the effect of attending college on voter turnout. Figure C1 below summarizes the effect estimates of the "scale of participatory acts" studies. The raw effect sizes (increase in number of acts performed) are not interpretably comparable to the effects on voter turnout (percentage-point change in propensity to vote). However, comparing within this group of studies yields a similar mixed picture. The studies by Kam & Palmer (2008) and Dinesen et al. (2016) do not detect a significant effect in any of their samples, and point estimates are quite small. However, confidence intervals indicate that they are, at the same time, unable to reject effect sizes of about 1 on a scale from 1–10. In other words, they are unable to reject effect sizes roughly equal to the significant estimates in Mayer (2011) and Henderson & Chatfield (2011), who reanalyze the data from the first study (#1) in Kam & Palmer (2008) using alternative procedures to perform the matching. In sum, the pattern in this body of evidence is mixed and undecided.

Figure C1. Effects of College on Scale of Participatory Acts.

Forest Plot summarizing existing evidence.



Note: Estimated effects of college education on political participation measured as a scale of participatory acts. Note that the studies by Mayer (2011) and Henderson & Chatfield (2011) use the same data as in the first (#1) study by Kam & Palmer (2008), but apply other kinds of cross-sectional matching.

C1.B. Additional Relevant Literature

A number of other studies – that do not specifically study the causal effect of interest in the current study (and in the studies in Figure 1) – have also been very important in informing the design, research question, and contribution of the current study. Firstly, a number of studies provide strong empirical evidence showing the degree to which different sets of confounders account for the general association between education and voter turnout. A study by Gidengil et al. (2019) uses data on siblings from Finland to show that family factors account for a large portion of the association between education and voter turnout. That is, they quantify, very convincingly, the extent to which selection on family characteristics affect this association. I use their findings to motivate the use of a within-person research design. In an innovative empirical study, Tenn (2005) shows that it is of crucial importance to take relative education in an individuals' cohort into account when estimating the effect of education. This stresses the importance of controlling for cohort effects, which the current study does by relying on two cohort panels. I do not include this latter study in Figure 1 because the study does not claim to provide an estimate of the causal effect of college. Specifically, Tenn (2005) finds evidence that their models' assumptions for causal inference are not met.

Secondly, a study by Tenn (2007) investigates an estimand that is close to the effect of attending college. The study convincingly estimates the marginal effect of one additional year of education by comparing schooling mainly *within* high school and college. It is important to note that this estimand is different since it does not investigate the margin *between* college and high school as the current study does. The analyses in Tenn (2007) yield a precisely estimated null effect. Thereby, it provides strong evidence that one more year of a given kind of education (e.g. high school or college) does not have any substantial effect on turnout. This suggests that if education has an effect, it is the difference between stages of education (i.e. upper secondary or tertiary education) that matters, not the continuous length or years of education (Stubager 2008). The current study and the studies summarized in Figure 1 try to assess exactly this. Disregarding the differences in terms of estimand, the study by Tenn (2007) is both well powered and indicates a null finding of an extra year of education. In the overall education–participation literature, it therefore goes against the pattern of statistically uncertain positive effects that I otherwise identify (cf. Figure 1). However, this just underlines the need for the current paper because we cannot know the answer to the following question: Is the null finding in Tenn (2007) mainly a consequence of comparing people who are at the same stage of education, or does it indicates more broadly that college education does not have a causal effect on voter turnout? By zooming in on the margin between college vs. no college, the current study provides evidence supporting the former answer.

C1.C. Studies on Political Knowledge

Studies of the effect of education on outcomes that are related to political participation also have important implications for the current study. Specifically, the family of studies on the effect on political knowledge has also been very convincing in proving the degree of selection into education by using both twin fixed effects and within-person designs. Firstly, studies by Weinschenk and colleagues find that controlling for twin fixed effects effectively neutralizes the effect of education found in less strict models (Weinschenk et al. 2021; Weinschenk and Dawes 2019). In a panel study, Highton (2009) provides credible evidence that differences in political sophistication between college-goers and non-college-goers were already in place before anyone attended college. Both of these studies show that unobserved characteristics – such as family socialization and genetics – both account for substantial selection into college education and have important consequences for political behavior. This further stresses the need for accounting for this in the current study, e.g. by using individual-level fixed effects.

C1.C. Studies on the Causes of Turnout

Finally, studies of non-educational causes of turnout more broadly have important implications for the current study. Specifically, they determine which alternative life events may explain or confound the relationship between education and voter turnout. Specifically, studies have shown the effects of factors such as health, residential mobility, labor market status, crime victimization, parenting, and marital status (Engelman et al. 2021; Highton and Wolfinger 2001; J.H. Hansen 2016; Sønderskov et al. 2020; Dahlgaard and Hansen 2021; Burden, Fletcher, Herd, Jones, et al. 2017). Therefore, I briefly review these studies in Appendix D, where I provide the reasoning for including specific time-varying covariates.

C2. Data and Research Design of Previous Studies

This appendix outlines the research designs and data used in previous studies of the effect of college on political participation.[2] The following appendix, C3, focuses on summarizing the statistical power of this body of research.

The study by Berinsky & Lenz (2011) reach an effect estimate with high internal validity by utilizing the Vietnam draft as an instrumental variable for attending college. They make a strong case for living up to the assumptions underlying the IV design. They thereby provide a great contribution to the literature and advance causal inference on this research question. While this is the case, their study leaves room for future studies due to three shortcomings that do not concern internal validity. The first is that their effect estimator is a local average treatment effect, i.e. they estimate the effect for compliers: those who only attend college to avoid the draft. The second shortcoming is that many things about going to college may have changed since the data was collected; the instrumentally "treated" group are men born in 1946–1948, i.e. attending college in the late 1960s and early 1970s. It is not straightforward that these circumstances may be generalized to more recent college-attending cohorts – in fact, we may expect college-goers today to be more prone to a mobilizing college effect (Burden 2009; Stevens, Armstrong, and Arum 2008).[3] Third, and arguably most important, is the high degree of statistical uncertainty in the instrumental variable analyses (see Appendix C3).[4]

The study by Dee (2004) uses geographical distance to a higher education institution as an instrument for attending college. The study generally concludes that education, including college, has an effect on voter turnout. The estimates for college are large but have high statistical uncertainty (see Appendix C3). The instrument used by Dee (2004) has been criticized since it may be that geographical distance to colleges are confounded with other factors such as income and parental education that may influence participation (Persson 2015).

The study by Apfeld (2021) employs a regression discontinuity design (fuzzy RD) with high causal credibility. They compare Romanians who score just above or below the cutoff grade required for entering university and find a large and statistically significant effect on a composite index of social capital, including voter turnout. They also report an estimate of the effect on dichotomous voter turnout in itself, which is applicable to the current study and is

[2] Many of these studies are discussed in a recent review by Persson (2015).
[3] I elaborate on this below.
[4] The point estimate in Berinsky and Lenz (2011) is 6 percentage points, and this design provides them with a standard error of 15 percentage points, yielding a very wide confidence interval and enabling them only to detect effects of 37 percentage points or higher with 70% power (Bloom 1995).

therefore the one I report in Figure 1 of the article. With their pre-registered innovative research design, the study in general constitutes a great advancement to the literature. The study's limitations are (a) the high statistical imprecision, (b) the local nature of estimate (pertains to those just around the cutoff grade), and (c) the potentially limited generalizability of the Romanian context. However, the study has high internal validity and makes a strong case for the Romanian context not being a deviant case.

The study by Dinesen et al. (2016) uses three different samples of twins from the US, Sweden, and Denmark. They employ a discordant twin design (twin fixed effects), which has clear advantages in ruling out selection bias due to genetic and family factors. Furthermore, the study provides a great advancement to the literature in terms of external validity by providing data on three different countries and different cohorts. They find insignificant estimates of the effect of college in all three samples but do find significant estimates for "years of education" and high school effects. Notably, the estimated effects for college are relatively more statistically uncertain compared to those for years-of-education and high school: In the US data, the effect estimate of college is numerically more than three times larger than the effect of one additional year of education. However, due to different standard errors, the effect of one additional year of education is statistically significant while that of college is not. Additionally, the effect estimate of college in the US is just as large as that of high school in Denmark. However, the latter is statistically significant in contrast to the former. Based on this, it may be argued that the basis for accepting an effect of years of education in the US and high school in Denmark is better than the basis for rejecting an effect of college in the US. Importantly, however, the null estimates of a college effect among Danish and Swedish twins are both more precisely estimated and numerically small.

The study by Burden et al. (2020) uses US data on siblings to estimate the effect of college on voter turnout. Importantly, beyond controlling for factors that are common to siblings, by using sibling fixed effects such as family environment, the study also controls for factors that differ *between* siblings that may affect both education and turnout. These include pre-adult measures of IQ and Big-5 personality traits. Another strength is that the study uses validated voter turnout as a dependent variable. The study complements the current study well since it estimates the longer-term effects of college on turnout and finds smaller effects that, with fixed effects included, are statistically insignificant or of borderline statistical significance. Notably, the study has the lowest degree of statistical uncertainty of the reviewed studies. The study does, nevertheless, underline the need for the current study for two reasons. Firstly, the study writes that the modest estimates may reflect a decaying education effect, due to the long-term nature of the estimated effects: They estimate the degree to which college affects participation that occurs "decades later." The theory of voting inertia does indeed suggest that the effect of education is largest in early adulthood before citizens settle into a habit of voting (Plutzer 2002). This is also in line with the general notion of formative or impressionable years in early adulthood, followed by stable political behavior (Kiley and Vaisey 2020; Sears and Funk 1999). Secondly, the siblings design may produce conservative estimates due to spillovers – although plausibly not to the same degree as twin designs – where the higher education of one sibling affects the less educated sibling (Dinesen et al. 2016; Gidengil et al. 2019). By using a different research design and synthesizing the body of prior studies, the current study addresses these questions and thereby advances the literature.

Finally, four studies have investigated the relationship in question via comparing the turnout levels of college-goers to a select group of non-college-attendees using matching techniques (Henderson & Chatfield 2011; Kam & Palmer 2008; Mayer 2011; Persson 2014). First, Kam & Palmer (2008) used propensity score matching on two US datasets to estimate the average treatment effect on the treated (ATT) of attending college. They found no effect on a composite index of participatory acts. They were later criticized by Mayer (2011) and Henderson & Chatfield (2011) who reanalyzed one of the datasets. The latter study concluded that the estimand simply may not be obtainable using matching. Both studies apply alternative matching methods and reported statistically significant and positive effect estimates on this basis. In sum, these studies leave a mixed picture. Persson (2014) used a cohort study from the UK to revisit the relationship using genetic matching. The study concludes against an effect of college based on a positive but statistically insignificant point estimate. In the current article, I present a difference-in-differences analysis based on a matching estimator. This estimation constitutes an improvement to previous matching studies in two ways. First, it improves on the statistical precision of all the previous matching studies. Second, I combine the causal leverage of matching with that of the difference-in-differences design. In other words, this design uses matching to try to obtain parallel trends in the potential outcomes rather than equal levels of potential outcomes.

More generally, it is relevant to note that most of the studies reviewed above primarily utilize data on citizens who attended college in the 1980s or earlier (Berinsky & Lenz 2011; Dee 2004; Dinesen et al. 2016; Henderson & Chatfield 2011; Kam & Palmer 2008; Mayer 2011; Burden et al. 2020). The exceptions are the UK data in Persson (2014), the Danish substudy in Dinesen et al. (2016), and the article by Apfeld et al. (2021) that uses data from Romania. Persson (2014) uses data on citizens attending college around 1990. The Danish subsample in Dinesen et al. (2016) went to college around 2000. The study by Apfeld et al. (2021) uses data on Romanians who went to university after 2015. As argued in the article, relying on older data means that respondents in previous studies are more likely to represent a more priviledged stratum of society. Importantly, this is a group that, due to the compensation hypothesis, may be less prone to the civic effects of education (Burden 2009; Stevens et al. 2008). A related characteristic is that the effect estimates of most existing studies – in order to obtain a high degree of internal validity – only pertain to a very specific (local) subgroup of college students, e.g. those who respond to a particular instrumental variable or fall close to a grade-point cutoff in an RDD. Thereby, they do not allow us to representatively conclude on the average effect of college on college-goers (Apfeld et al. 2021; Berinsky & Lenz 2011; Dee 2004). The current study makes large improvements on this by estimating the ATT for all college-goers in a nationally representative panel of US voters.

C3. Minimum Detectable Effect Sizes of Prior Studies of College and Voter Turnout

In Table C3, I report – as a supplement to Figure 1 in the article – the minimum detectable effect sizes (MDE) of the studies. These are calculated based on the rules in Bloom (1995) to obtain the MDE for a two-sided test at 70–80% power and 95% confidence level. Note that the statistical imprecision is especially large ($MDE > 24$ percentage points) in the three studies that use a natural experiment or RDD (Apfeld et al. 2021; Berinsky & Lenz 2011; Dee 2004) compared to the matching study by Persson (2014) and the sibling study by (Burden et al. 2020). For example, Berinsky and Lenz (2011) reach an estimate with high internal validity by utilizing the Vietnam draft as an instrumental variable. However, this design provides them with a standard error of 15 percentage points, yielding a very wide confidence interval and enabling them only to detect effects of 37 percentage points or higher with 70% power.

Table C3 *Minimum Detectable Effect Sizes in Studies of College and Voter Turnout*

Study	**MDE (80% power)**	**MDE (70% power)**	**Study Design**
Burden et al. 2020 [USA]	5.6	5.0	*Sibling FEs*
Persson 2014 [United Kingdom]	8.5	7.5	*Genetic Matching*
Berinsky & Lenz 2011 [USA]	42.0	37.2	*Vietnam Draft IV*
Dee 2004 [USA]	27.1	24.0	*College Distance IV*
Apfeld et al. 2021 [Romania]	27.7	24.6	*Grade Cutoff RDD*
This Study [USA] ***Full Sample***	**4.5**	**4.0**	*Individual-Level DiD*
This Study [USA] ***Restricted***	**7.8**	**6.9**	*Individual-Level DiD*
This Study, Replication Sample [USA] *Full Sample*	3.4	3.0	*Individual-Level DiD*
This Study, Replication Sample [USA] *Restricted*	6.7	5.9	*Individual-Level DiD*

Note: MDE denotes the Minimum Detectable Effect Size for a two-sided test with 95% confidence level and 70–80% statistical power (cf. columns). The MDE is "the smallest effect that, if true, has a [70 or 80]% chance of producing an impact estimate that is statistically significant at the [95]% level" (Bloom 1995). Calculations are based on standard errors of existing studies and formulas from Bloom (1995).

C4. Meta-Analysis

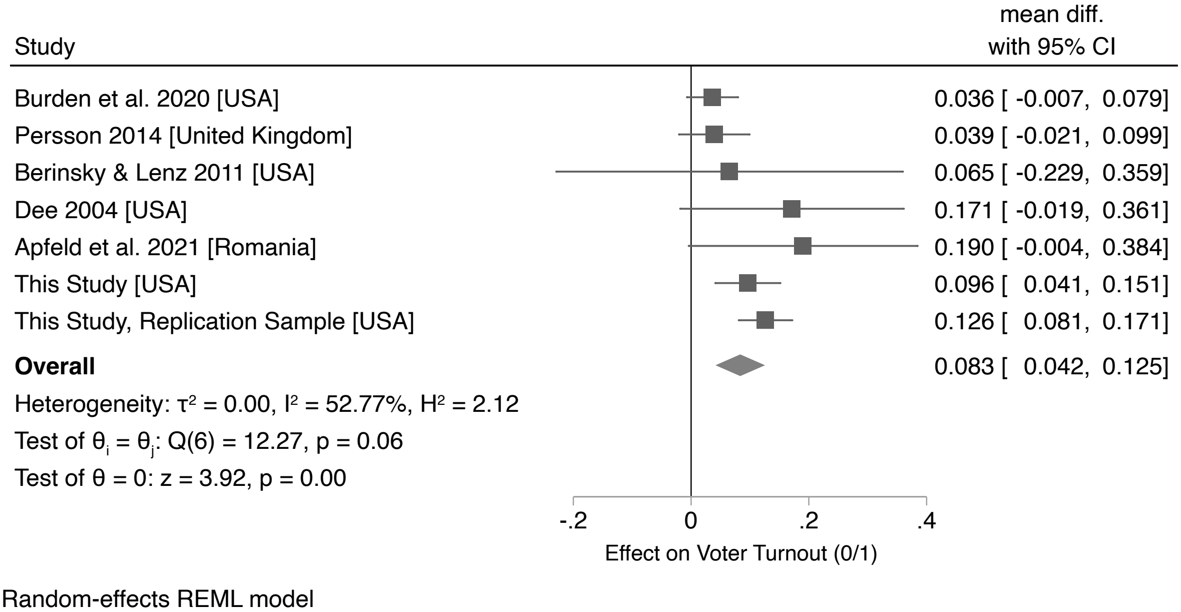
This appendix provides the methodological details and robustness checks of the meta-analysis presented in Figure 1 of the main article. The included studies are described and discussed in depth in Appendix C2. Each of the studies in Figure 1 estimate the effect of attending some college on a binary turnout variable. This contributes to the comparability of the studies in terms of their harmonization (Slough and Tyson 2023). However, as mentioned in main article, the individual studies draw on quite different national contexts, subpopulations and leverage different sources of variation in education. Importantly, the aim of this analysis is to view prior causally credible studies of the college effect in conjunction and test their joint compatability with a positive effect rather than to meta-analyze studies that are exactly the same on all dimensions (Holbein et al. 2023). The meta-analysis relies on point estimates and standard errors as reported by the original study authors.

Figure C2 presents the detailed results of the meta-analysis presented in Figure 1 of the main article. The figure presents the effect sizes and confidence intervals of each study and the meta-analytic average (See Figure C4 for the meta-analytic average using only prior studies). This specification uses a random effects model and studies are weighted by inverse variance. Figure C3 presents the highly similar results obtained using a fixed effects model (DerSimonian and Laird 1986).

Finally, I conduct a series of robustness checks of the meta-analytical findings. Across specifications and excluding various sets of studies, I find that prior studies are compatible with a positive effect of college. First, all robustness checks below leave out the replication sample from the current study to make estimates conservative. Second, Figure C4 presents a meta-analysis that leaves out the estimates obtained from the current study's difference-in-differences analyses. Third, Figure C5 presents a meta-analysis that only relies on studies from the US context. Fourth, Figure C6 distinguishes between studies that estimate short-term versus longer-term effects of college.

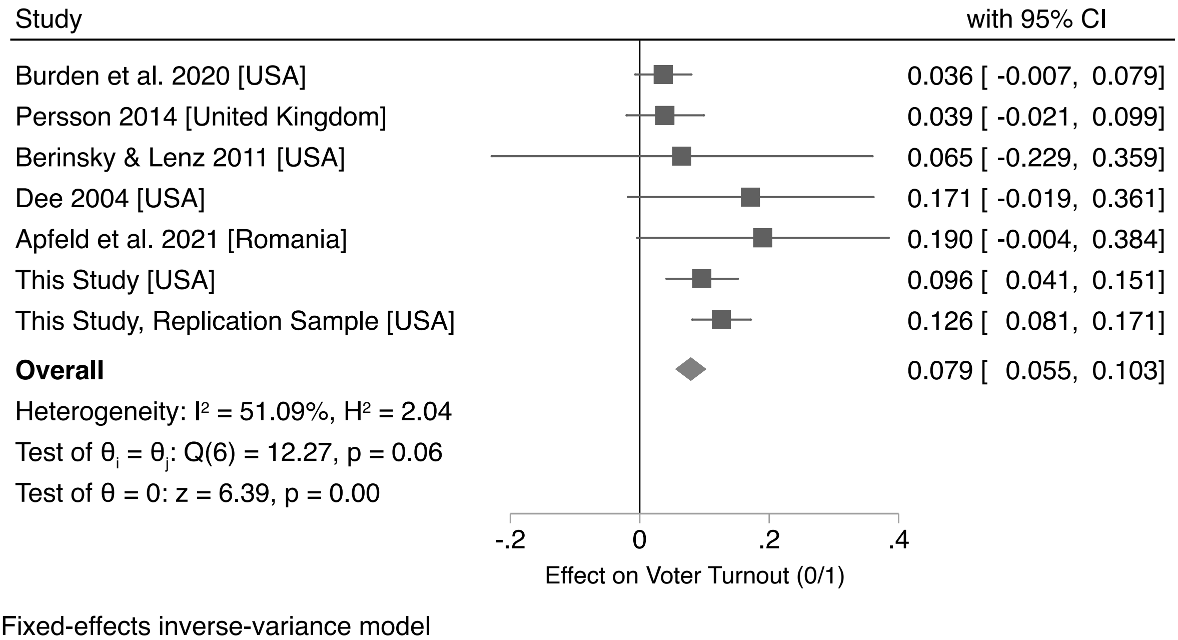
The meta-analyses were performed using the `meta` package in Stata.

Figure C2 *Meta Analysis*



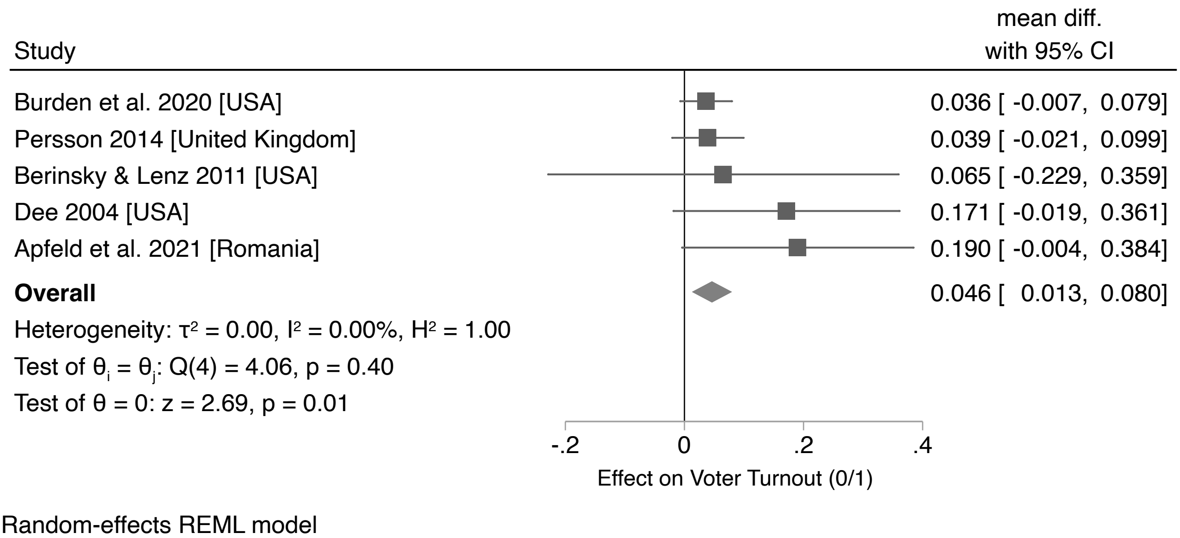
Note: Results from a formal meta-analysis using random effects weighting studies by inverse variance. Grey triangles represent the 95% confidence intervals for a meta-analytic average of effect sizes.

Figure C3 *Fixed effects*



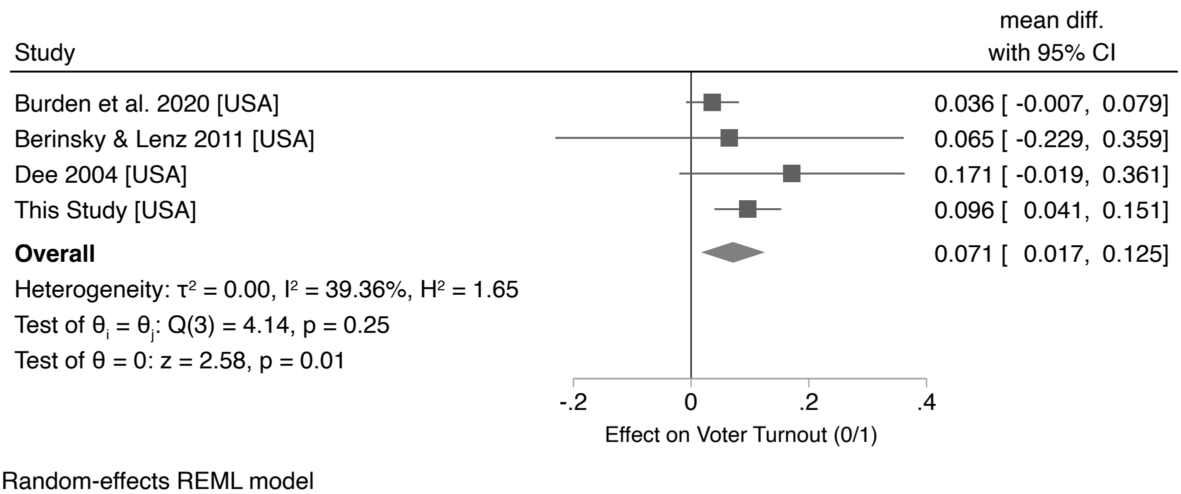
Note: Results from a formal meta-analysis using random effects weighting studies by inverse variance. Grey triangles represent the 95% confidence intervals for a meta-analytic average of effect sizes.

Figure C4 *Excluding this study's estimates*



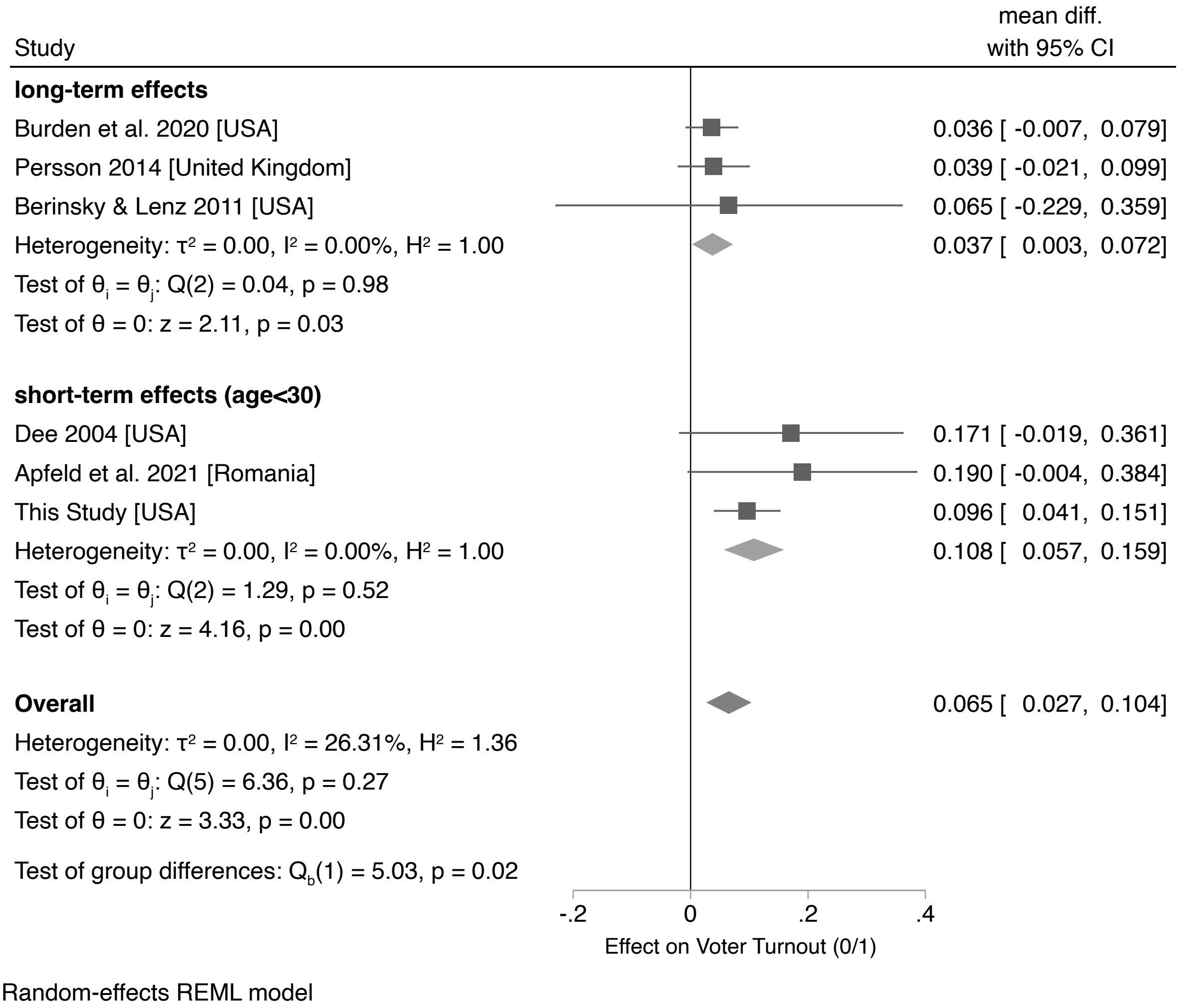
Note: Results from a formal meta-analysis using random effects weighting studies by inverse variance. Grey triangles represent the 95% confidence intervals for a meta-analytic average of effect sizes.

Figure C5 *Including only studies from the US context*



Note: Results from a formal meta-analysis using random effects weighting studies by inverse variance. Grey triangles represent the 95% confidence intervals for a meta-analytic average of effect sizes.

Figure C6 *Subgrouping studies by estimand: short vs. long-term effects of college*



Note: Results from a formal meta-analysis using random effects weighting studies by inverse variance. Grey triangles represent the 95% confidence intervals for a meta-analytic average of effect sizes.

Appendix D: Considerations about Covariates and Confounders

D1. Time-varying Covariates

In order for the estimates reported in this article to be causal, we rely on the parallel trends assumption (PTA). This assumption requires that the panelists who went to college would – had they not gone to college – have experienced the same development over time in voter turnout as those who did not go college. One of the things I do in order to meet this assumption is include a number of time-varying covariates in order to rule out that the changes in turnout, which we attribute to attending college, actually are due to these time-varying phenomena. I describe these variables in the article but elaborate here on the grounds for including them. First of all, health has been shown to be important in determining voter turnout (Engelman et al. 2021; Burden, Fletcher, Herd, Moynihan, et al. 2017). Geographically moving to a new community or whether you live with your parents or not is also shown to disrupt voting patterns (J.H. Hansen 2016; Highton and Wolfinger 2001). Being the victim of violent crime has also been shown to change peoples' voting participation, and fortunately, I am also able to control for this (Sønderskov et al. 2020). Finally, getting married, having kids, and losing your job may also be a source of differential turnout trends, and therefore, I include these variables in my set of time-varying controls (Highton and Wolfinger 2001; Dahlgaard and Hansen 2021).

D2. Adjusting Turnout Trends for Differences in Pre-College Turnout

Based on the literature on voter turnout, an important potential confounder is whether you have voted before (Coppock and Green 2016; Gerber, Green, and Shachar 2003; Plutzer 2002). In our case, if treated and untreated panelists are systematically different on pre-college turnout, it may be that their respective change or stability in turnout (the outcome of our difference-in-differences model) is due to these differences in pre-college participation rather than whether they attended college. To rule out that changes or lack of changes in turnout over time is due to the well-established habit-forming effect of voting (Coppock and Green 2016; Gerber, Green, and Shachar 2003; Plutzer 2002), I estimate the counterfactual trends for treated respondents using untreated respondents who were similar on pre-college turnout. In terms of estimation, this is done by interacting pre-college turnout with the time-fixed effects, which allows pre-college voting to have over-time[5] effects on turnout (Hainmueller and Hangartner

[5] I use the terms dynamic effects and over-time effects interchangably – the denote the effects that a variable that is stable over time may have on changes in a dependent variable over time.

2013; Hall and Yoder 2022; Schafer et al. 2022, see also appendix D5).[6] This means the effect estimates are obtained by comparing the turnout trend among treated respondents who voted in 2004 with the trend among untreated respondents who voted in 2004 (and similarly compares turnout trends among treated respondents who did not vote in 2004 with the trend among untreated respondents who did not vote in 2004).

In addition to the theoretical reasoning based on habit forming theory above, there is a methodological/mathematical reason to adjust trends for pre-college voting, which is due to ceiling and floor effects that are inherent to the dichotomous measurement of voter turnout. In the following, I will show that these ceiling and floor effects interact in a problematic way with pre-existing imbalances on turnout between treated and untreated respondents – except in the odd case where it is true that nobody changes their turnout over time or in the odd case where the bias induced by this imbalance is cancelled out by other confounding. The implication is that unless nobody changes their turnout over time, then an imbalance on pre-treatment turnout between treated and untreated respondents would cause the parallel trends assumption (PTA) to be violated. This is true even if no other factors influence trends in turnout over time. This provides another strong argument for – as described above – adjusting the estimates for over-time effects of such pre-treatment differences in voter turnout. Below, I will provide formal proof to support this idea.

Specifically, I seek to prove that in the case where we do not control for the time-varying effects of pre-treatment voting, then the following proposition holds:

Proposition A: If there is an imbalance on pre-treatment turnout between treated and untreated respondents, then the parallel trends assumption is violated.

[6] Note that interacting pre-college turnout (i.e. 2004-turnout) with the time-fixed effects is, in our specific case, mathematically equivalent to including "prior turnout" as a time-varying covariate. This is because that strictly speaking no one voted in the 2000 election (i.e. in the 2004 election "prior voting" is equal to zero for all observations, and in 2008 "prior voting" is equal to 2004 turnout). However, note that this is, in terms of post-treatment bias, a "safe" time-varying covariate – neither 2000 or 2004 voter turnout can be affected by the treatment (Rohrer 2018).

A simple version of this proof is based on a control variable logic and takes the following form:

Proof of Proposition A:

(1) Assume that pre-treatment turnout affects *changes* in turnout.

(2) Assume then that we do not control for dynamic effects of pre-treatment turnout.

(3) Note that the following is true: if a pre-treatment variable (Z), that we do not control for the dynamic effects of, is unevenly distributed between treated and untreated respondents **and** that variable affects changes in turnout, **then** the parallel trends assumption will be violated unless we control for the dynamic effects of that variable (Z).

   a. It follows that if we do not control dynamically for a variable like Z, then the parallel trends assumption is violated.

(4) Observe, finally, that if there is an imbalance on pre-treatment turnout between treated and untreated respondents **and** pre-treatment turnout affects changes in turnout over time (1), then it follows given (2) that pre-treatment turnout is variable like Z in (3).

(5) It follows, given (1) and (2), that if there is an imbalance on pre-treatment turnout, then the parallel trends assumption is violated. This is what we wanted to prove. □

Thus, all that remains is to prove what is assumed in statement (1): that pre-treatment turnout affects changes in turnout. Note that we may simply assume this to be the case based on prior studies that show voting to be habit forming – and then we would be done (Coppock and Green 2016; Gerber et al. 2003; Plutzer 2002). This effect on persistence in voting is exactly what is needed to satisfy the second condition for being a variable like Z in statement (3). However, if we do not wish to make this assumption, I provide a mathematical argument below. This is where the ceiling and floor effects come in. Note that all of the places that mention voting/turnout below are referring to potential outcomes in the absence of treatment. Further, let $voted_{pre}$ and $voted_{post}$ denote whether you voted in the pre-treatment period and the post-treatment period, respectively.

Proposition B: Pre-treatment turnout affects changes in turnout.

Proof of Proposition B:

(1) We know that turnout is dichotomous: You either vote or you do not vote (i.e. $voted \in \{0,1\}$).

(2) It follows that *changes* in turnout $(\Delta voted = voted_{post} - voted_{pre})$ may take three different values:
- Either you move from voting $(voted_{pre} = 1)$ to not voting $(voted_{post} = 0)$, which means that $\Delta voted = -1$ (i.e. change was negative).
- Or you move from not voting $(voted_{pre} = 0)$ to voting $(voted_{post} = 1)$, which means that $\Delta voted = 1$ (i.e. change was positive).
- Or you did not change your turnout $(voted_{pre} = voted_{post})$, which means that $\Delta voted = 0$ (i.e. change was zero).

(3) It follows that whether you voted in the pre-treatment election restricts the potential direction in which your turnout can change. In other words, pre-treatment turnout restricts the values that changes in turnout may take.
- Specifically, if you voted in the pre-treatment election $(voted_{pre} = 1)$, then your turnout may only be unchanged or decreased in the following election, i.e. either you vote again $(voted_{post} = 1)$, which means that change in turnout is zero $(\Delta voted = 0)$, or you do not vote again $(voted_{post} = 0)$, which means that change in turnout is negative $(\Delta voted = -1)$. *It follows that the change in turnout is restricted to being non-positive.*
- In the same way, if you did not vote in the pre-treatment election $(voted_{pre} = 0)$, then your turnout may only be unchanged or increased in the following election, i.e. either you stay a non-voter $(voted_{post} = 0)$, which means that change in turnout is zero $(\Delta voted = 0)$, or you change from not voting to voting $(voted_{post} = 1)$, which means that change in turnout is positive $(\Delta voted = 1)$. *It follows that the change in turnout is restricted to being non-negative.*

(4) It follows that only pre-treatment voters $(voted_{pre} = 1)$ may experience negative changes in turnout. Moreover, it follows that only pre-treatment non-voters $(voted_{pre} = 0)$ may experience positive changes in turnout.

(5) Unless nobody changes their turnout (i.e. $\Delta voted = 0$ for everyone), if we performed an ideal randomized experiment that randomly assigned respondents to either $(voted_{pre} = 1)$ or $(voted_{pre} = 0)$, we would necessarily find that the *change* in turnout was larger among the group with $voted_{pre} = 0$ than in the group with $voted_{pre} = 1$. More specifically, the average change in turnout among those with $voted_{pre} = 1$ would belong to the interval $[-1; 0[$. The average change in turnout among those with $voted_{pre} = 0$ would belong to the interval $]0; 1]$.

(6) It should be clear from (3), (4), and (5) that pre-treatment voting in the relevant sense affects changes in turnout. This is what we wanted to show. □

Note that I only prove Proposition B under the assumption that the following is false: "Proposition C: Nobody changes their turnout, i.e. change in turnout is zero for all respondents" (this is introduced in step (5) of the proof of Proposition B). This means that the proof of Proposition A is also only valid if Proposition C is indeed false. However, it seems evident that Proposition C being false is a very plausible assumption. Indeed, it can be confirmed by finding just one observation in a given dataset that changes from voting to not voting or from not voting to voting. Conversely, it would be highly unlikely to observe a dataset where all who voted in one election would also turn out in the second election and where all who did not vote in one election would also not vote in the second election. Therefore, this condition is mainly of theoretical importance and is not explicated everywhere in the above proofs.

As a final interesting note, it should be remarked that the randomized experiment in step (5) of the proof of Proposition B is indeed **not** necessarily expected to find a difference in post-treatment turnout *levels* (i.e. an effect on the average values of $voted_{post}$), but only in the post-treatment *changes* in turnout (i.e. on the average values of $\Delta voted$. At least, this does not follow from the proof. Such an effect on levels is not necessary for the proofs to hold.

D3. Regressions without Covariates

This section reports regression results with different combinations of covariates included. We may distinguish between covariates that may possibly be caused by the treatment (and introduce post-treatment bias) and those covariates that may not be caused by the treatment (Elwert and Winship 2014; Montgomery, Nyhan, and Torres 2018). Note that the fixed covariate of pre-treatment turnout may not be affected by the treatment (whether you voted at the election before going to college may not be caused by attending college). The Pre-college Voting × Time fixed effect described in section D2 is thereby a "safe" covariate in this sense. This is not true about the time-varying covariates outlined in section D1, namely residential mobility, living with parents, marital status, crime victimization, becoming seriously ill or disabled, job loss, and having children. Due to this – as well as the considerations and proof provided in Appendix D2 above – the main article reports models both with and without the "unsafe" covariates, while all models include the Pre-college Voting × Time Fixed Effects. However, for reference, I provide the regressions with the remaining possible combinations of covariates below (Table D1). Note that the two groups (college-attendees and non-attendees) defined by the main independent variable of this study do indeed differ substantially on pre-college (2004) turnout: Table F3 in Appendix F shows that 2004 non-voters constitute a majority in the untreated group but a minority in the college-going group. In light of the above arguments, this means that we should expect the estimates to be different when we (do not) adjust turnout trends for this imbalance. Specifically, the models that do not take the arguments made above into account by allowing pre-college turnout to have dynamic effects are models 3, 4, 7, and 8. The preferred models (1, 2, 5, and 6) instead compare the turnout trend among treated respondents who voted in 2004 with the trend among untreated respondents who voted in 2004 (and similarly compares turnout trends among treated respondents who did *not* vote in 2004 with the trend among untreated respondents who did *not* vote in 2004). The table shows that, if we do not adjust for dynamic effects of pre-college voting, then the effect estimate changes in a negative direction. This is consistent with the arguments made above. Finally, we may use the placebo analysis performed in appendix A2 to assess the appropriateness of adjusting trends for pre-college differences in turnout. Note that the placebo analysis tests whether future college-goers differ in their propensity to experience increases or decreases in turnout compared to future non-attendees. The fact that the placebo-estimates are small and insignificant – even though we include Pre-college Voting × Time FEs – suggests that the substantial and statistically significant effect estimates in the main analyses are not an artefact of this adjustment procedure.

TABLE D1 *Regressions with different combinations of time-varying covariates*

	Full Sample				Restricted Ctrl.grp			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No College	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Attended College	0.130***	0.137***	-0.010	-0.020	0.106***	0.096***	0.011	-0.030
	(0.016)	(0.014)	(0.018)	(0.016)	(0.031)	(0.028)	(0.037)	(0.034)
Time FEs	✓	✓	✓	✓	✓	✓	✓	✓
Individual FEs	✓	✓	✓	✓	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓			✓	✓		
Time-varying Controls	✓		✓		✓		✓	
Units	9,652	10,426	9,667	11,131	8,717	9,360	8,731	9,902
Observations	18,298	19,813	18,313	20,518	16,593	17,842	16,607	18,384

Note: Difference-in-differences estimates of how college-goers changed their turnout compared to respondents who did not go to college between the 2004 and 2008 elections. In models 5 through 8 the untreated group is restricted to respondents who went to college after 2008 (future college-goers). Robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. Models 1,2 5 and 6 include time-trends that are specific to whether a respondent voted in the pre-treatment election or not. Note that the models (3, 4, 7 and 8) that do not adjust for the dynamic effects of pre-treatment voting yield a null estimate, which is likely due to the type of bias discussed and proven in Appendix D above.

D4. Bundled Treatment and Time-Varying Covariates

Another challenge to identification in the applied difference-in-differences design concerns the issue of bundled treatment. That attending college is likely correlated with other changes that may affect turnout means that our design estimates the combined effect of attending college along with such correlated changes. However, the fact that our estimates are robust to including the life events we measure here (e.g. controlling for moving to a new place) constitutes tentative evidence that the estimated effect of college is not driven by these correlated mechanisms.

Appendix E: Restricted Comparison Group and Additional Robustness Checks

One of the identification strategies applied in the article utilizes the group of late-attendees as a counterfactual group (i.e. models 4–6 in Table 1 of the main article). These models are estimated completely in the same way as the full sample models based on equation (1) of the main article, except that those respondents who both did not go to college from 2004 to 2008 and did not attend college *after* 2008 are omitted from the sample. This means that the comparison group is restricted to respondents who did not go to college from 2004 to 2008 but who did attend college after 2008 (specifically between 2008 and 2012). The identifying assumption is that in terms of trends in turnout, late college-goers are similar to early college-goers. Below, I further probe this assumption.

In the first three models of Table E1, I add further restrictions to the restricted comparison group models. In the first and second columns of Table E1, I restrict the groups to be comparable on military service by dropping respondents who served in the military. In the third column, I add military service as a time-varying covariate. This is intended to address the potential issue that military service may confound the relationship by being correlated with both attending/delaying college and voting at elections. This concern may also be applied to the full sample models. Therefore, in columns 4–6 of table E1, I perform the same analyses using the entire sample (i.e. restricting to non-military service respondents in models 4–5 and adding military service as time-varying covariate in model 6). In all models, I find that coefficients are almost identical to those estimated in the main table of the article. These results indicate that differences in the frequency and timing of military service are not driving the main results.

Another potential concern with using the future attendees as comparison group is whether the differential timing of going to college may be driven by factors that also influence changes in turnout. To make the timing of college as similar as possible we may, therefore, further restrict the comparison group to those who attend college *just* after the 2008 November election. Table E2 presents the difference-in-differences results based on different restrictions on the timing of future college attendance. Specifically, model 1 shows the results with the untreated group restricted to only those who start college between one month and one year after the 2008 election. The differential timing of college attendance is arguably more ignorable for this group than for those who start college up to three years later. Model 2 adds another year to this period, such that the untreated group consists of those starting college in either 2009 or 2010. Models 3 and 4 each add one additional year. The results are remarkably robust to these alternate specifications, as we see the effect estimates being stable across models with only the statistical uncertainty varying. As expected, the more restrictive models have larger standard errors. This suggests that factors related to the timing of college does not confound the results from the future attendees models.

Table E1 *Robustness to variations in military service*

	Restr. Ctrl.grp w. military dropped		Restr. Ctrl.grp	Full Sample w. military dropped		Full Sample
	(1)	(2)	(3)	(4)	(5)	(6)
No College	ref.	ref.	ref.	ref.	ref.	ref.
Attended College	0.095**	0.108***	0.104***	0.133***	0.125***	0.127***
	(0.029)	(0.031)	(0.030)	(0.015)	(0.016)	(0.015)
Time FEs	✓	✓	✓	✓	✓	✓
Individual FEs	✓	✓	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓	✓	✓
Time-varying Controls		✓	✓		✓	✓
Units	9,233	8,750	8,862	10,269	9,675	9,815
Observations	17,589	16,702	16,925	19,501	18,414	18,692

Note: Difference-in-differences estimates of how college-goers changed their turnout compared to respondents who did not go to college between the 2004 and 2008 elections. In models 1-3, the untreated group is restricted to respondents who went to college after 2008 (future college-goers). In models 1, 2, 4 and 5 it is further restricted to those without military service. Robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All models include time trends that are specific to whether a respondent voted in the pre-college election or not. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. In models 3 and 6 it also includes Military Service.

Table E2 *Further restricting the timing of future college attendance*

	fut. attend. ≤ 2009	fut. attend. ≤ 2010	fut. attend. ≤ 2011	fut. attend. ≤ 2012
	(1)	(2)	(3)	(4)
No College	ref.	ref.	ref.	ref.
Attended College	0.095*	0.090*	0.106***	0.096***
	(0.047)	(0.036)	(0.031)	(0.028)
Time FEs	✓	✓	✓	✓
Individual FEs	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓
Time-varying Controls				
Units	9,145	9,239	9,315	9,360
Observations	17,439	17,614	17,757	17,842

Note: Difference-in-differences estimates of how college-goers changed their turnout compared to respondent who did not go to college between the 2004 and 2008 elections. The untreated group is restricted to respondent who went to college after 2008 (future college-goers). Model 1 is restricted to future college-goers who g to college *just* after the 2008 election (i.e. starting no later than 2009). Models 2-4 loosen this restriction b including an additional year each (i.e. starting no later than 2010, 2011 and 2012, respectively). Robust standar errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All models include tim trends that are specific to whether a respondent voted in the pre-college election or not. Time-varying control include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill o disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the famil parent died and relative or friend died.

Appendix F: Details and Descriptive Statistics on the ELS:2002

F1. Detailed Information about the ELS:2002 Sample

This study uses the publicly available data from the Education Longitudinal Study of 2002 (ELS:2002). The ELS:2002 was collected by the National Center of Education Statistics of the U.S. Department of Education as part of a program of several longitudinal cohort studies (Ingels et al. 2014; NCES 2015). The study comprises a survey component and an administrative transcript component. The survey component consists of four waves collected in 2002, 2004, 2006, and 2012, respectively. Initially, a nationally representative sample of the cohort of high school sophomores in the spring term of 2002 were sampled. The initial wave sample was based on a two-stage, stratified probability design. In the first stage, high schools were chosen. In the second stage, sophomore students were selected from these schools. In the first follow up (2004), the sample was freshened to be representative of the cohort of high school seniors in the spring term of 2004. This freshened sample comprised 16,763 sample members (16,525 eligible from the base-year collection; 238 identified as part of sample freshening). The current study is based on respondents from this senior cohort who also answered the third and fourth waves of the survey (2006 and 2012). The third and fourth waves were conducted two and eight years after scheduled high school graduation, respectively. Finally, the administrative transcript component was collected in 2013 and consists of postsecondary education transcripts. Table F1 shows the response rates for the different survey waves (Table 1 in Christopher 2014). I rely on data from all waves of the survey, which gives a retention rate of 77.8%.

Table F1 *Eligible Sample and Weighted Response Rates for ELS:2002*
Source: Table 1 of Christopher (2014)

Study round	Eligible[1]	Weighted response rate
Base-year	17,754	87.4
First follow-up	16,733	88.7
Second follow-up	16,700	81.9
Third follow-up	16,562	77.8
Postsecondary transcripts	12,549	77.2

[1] Eligibility is based on membership in the sophomore or freshened senior cohort. Detected sampling errors (based on erroneous school report of grade level)—and deceased individuals—have been removed. In addition, sample members who did not report postsecondary education participation have been removed from the postsecondary transcript response rate denominator.

F2. Panel Attrition and Representativeness

Below, Appendix F2.A analyzes sample attrition. Appendix F2.B examines the degree to which the underrepresentation of college non-attendees in the sample may affect the external and internal validity of the findings.

F2.A. Panel Attrition

As Table F1 above indicates, the survey response rate fell between the first wave of the survey and the last wave of the survey. In the following, I examine whether this dropout poses potential problems to the results reported in the main analyses.

First, to examine whether the change in the proportion of missing cases is systematic, I investigate which first-wave characteristics predicted dropping out of the survey. Table F2A1 reports estimates from a linear probability model. The sample is initially restricted to wave one respondents, and the dependent variable is then whether the respondent was missing in the final wave on which this study relies. We see that respondents from homes with a higher socio-

TABLE F2A1 *Predicting Sample Attrition (dropout) Between Survey Waves*

	Dependent Variable: Dropped Out of Survey (0/1)	
Race: Hispanic	ref.	
Race: Other	0.014	(0.014)
Race: Asian	-0.001	(0.016)
Race: White, Non-Hispanic	-0.052***	(0.012)
Sex: Female	ref.	
Sex: Male	0.092***	(0.007)
Parental SES 1st quartile	ref.	
Parental SES 2nd quartile	-0.034**	(0.011)
Parental SES 3rd quartile	-0.054***	(0.011)
Parental SES 4th quartile	-0.108***	(0.011)
Constant	0.284***	(0.012)
N	13469	

Note: Linear Probability Model predicting dropping out between the first wave of the survey and the last wave used in this study. Robust standard error sin parentheses. * p<0.05, ** p<0.01, *** p<0.001.

economic status are less likely to drop out. This is also true for white, non-Hispanic respondents. Finally, male respondents are substantially more likely to drop out of the survey. Results are the same using a logit model.

Secondly, to make the implications for internal validity more tangible, I correlate attrition with the propensity score for going to college. Table F2.A2 reports this association estimated with a linear probability model (results are virtually identical with logit models). Unsurprisingly, those more likely to attend college are less likely to drop out and stop responding to the survey. This is true when using either a continuous or a decile measure of the propensity score for going to college. In terms of voting, we also know that that people who do not respond to surveys are less likely to turn out to vote (Selb and Munzert 2013; Sciarini and Goldberg 2016; Burden 2000). This means that we are more likely to overestimate the turnout in 2008 (and thereby the growth in turnout) among non-college-goers, than among college-goers. This has implications for the causal estimates of the main analyses as it suggests that the estimated effect of going to college is actually attenuated by this attrition-related mechanism.

Thirdly, to further examine this, I estimated respondents' propensity to drop out of the survey, and then ran the main analyses conditional on this propensity to drop out. By analyzing

TABLE F2A2 *Predicting Sample Attrition (dropout) Between Survey Waves Using the Propensity Score for Attending College*

	Dependent Variable: Dropped Out of Survey (0/1)	
	(1)	(2)
Propensity Score (continous)	-0.397*** (0.025)	
Prop.Score: 1. decile		ref.
Prop.Score: 2. decile		-0.017 (0.018)
Prop.Score: 3. decile		-0.059*** (0.016)
Prop.Score: 4. decile		-0.077*** (0.017)
Prop.Score: 5. decile		-0.092*** (0.018)
Prop.Score: 6. decile		-0.128*** (0.016)
Prop.Score: 7. decile		-0.115*** (0.016)
Prop.Score: 8. decile		-0.152*** (0.015)
Prop.Score: 9. decile		-0.197*** (0.015)
Prop.Score: 10. decile		-0.152*** (0.023)
Constant	0.572*** (0.021)	0.345*** (0.012)
N	13469	13469

Note: Linear Probability Models predicting dropping out between the first wave of the survey and the last wave used in this study. Propensity scores were calculated using the supplementary matching method described in appendix I3. Model 1 tests a linear assocation, whereas Model 2 breaks the propensity score ito deciles. Robust standard error sin parentheses. * p<0.05, ** p<0.01, *** p<0.001.

whether the effect of college varies between groups that are more vs. less likely to drop out of the survey, we are able to shed some light on whether the estimated effects generalize to these dropouts. The underlying assumption is that we can use those respondents who had a high propensity to drop out but did not drop out to learn about those who did actually drop out. Table F2.A3 reports estimates from an interaction model where equation (1) of the main article has been extended with a full set of interactions between the treatment variable, the time fixed effects, and the estimated propensity to drop out of the survey. If the interaction coefficient is positive, it means that the effect is larger among those with high propensity to drop out. We see that the difference in effect size between those with different propensities to drop out of the survey is statistically indistinguishable from zero. Using a categorical indicator for high vs. low propensity scores yields a rather precise null estimate of the interaction. The statistical uncertainty is remarkably larger using a continuous measure of propensity scores. Robustness checks suggest, however, that assuming such a linear model does not seem to be tenable (Hainmueller, Mummolo, and Xu 2019). This evidence suggests that the effect of college does

TABLE F2A3 *Does the Estimated Effect Of College Vary Between Respondents With High vs. Low Propensity to Drop Out Of The Survey?*

	Categorical Propensity Score (High/Low)			Continous Propensity Score		
	(1)	(2)	(3)	(4)	(5)	(6)
No College	ref.	ref.	ref.	ref.	ref.	ref.
Attended College	0.121***	0.112***	0.098***	0.159**	0.133*	0.106
	(0.023)	(0.025)	(0.026)	(0.055)	(0.060)	(0.063)
Attended College × Low Propensity Score (cat.)	ref.	ref.	ref.			
Attended College × High Propensity Score (cat.)	0.001	0.008	0.011			
	(0.029)	(0.032)	(0.032)			
Attended College × Propensity Score (cont.)				-0.138	-0.059	-0.003
				(0.197)	(0.215)	(0.223)
Time FEs & Individual FEs	✓	✓	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓	✓	✓
Time-varying Controls		✓	✓		✓	✓
Additional Pre-college Covariates × Time FEs			✓			✓
Units	10,422	9,652	9,554	10,422	9,652	9,554
Observations	19,806	18,298	18,118	19,806	18,298	18,118

Note: Interaction model testing the difference in effect size between those who had high vs. low propensity to drop out of the survey. If the estimated interaction is positive, it means that the effect is larger among those with higher propensity to drop out. Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. All models include time trends that are specific to whether a respondent voted in the pre-college election or not. Models 3 and 6 further interacts time fixed effects with pre-college cognitive skills, gender, race and parental education and income. Robust standard errors clustered by individual in parentheses.* p<0.05, ** p<0.01, *** p<0.001.

not vary with propensity to drop out of the survey. This increases confidence in both the internal and external validity of the estimated causal effect of college on college-goers: The absence of those who dropped out seemed to pull the effect neither up nor down.

F2.B. Representativeness

The analytical sample, after accounting for longitudinal survey attrition and missing data on relevant variables, disproportionately consists of college-goers compared to non-college-goers. As mentioned in the main article, 87% of the analytical sample are college-goers. The respondents in the ELS:2002 were primarily born in 1986, 1985 and 1984. For comparison, data from the CPS shows that 43% of those born between 1984 and 1988[7], had obtained an associate degree or higher in 2013. This means that we are, to a much lesser extent, reaching the non-college-goers in the cohort of interest with our sample.[8] This fact may be considered both in terms of its implications for external validity and internal validity. Firstly, regarding external validity, it is important to recall what the estimand of the current study is: We are trying to isolate the so-called Average Treatment Effect on the Treated (ATT) – i.e. the effect that attending college has on those people who go to college. This is opposed to the Average Treatment Effect (ATE), where we are also interested in the effect that attending college would have on those in our population who do not attend college. Since we are only interested in the former, and thereby are not interested in generalizing our findings to the population of non-college-goers, it is not problematic that our survey does not adequately represent non-college-goers. Thus, in terms of external validity of the estimated ATT, the overrepresentation of college-goers does not, in itself, pose a problem. Note, however, that if it was the other way around – that college-goers were underrepresented in our sample – we should be more concerned about the representativeness of our sample for the population of interest. Finally, it is important to note that there is also non-response in the college-going group, although not to the same extent as among non-college-goers. The representativeness of our findings does, therefore, mainly extend to the kind of college-goers who reply to surveys.

Secondly, the underrepresentation of non-college-goers may have implications regarding the internal validity of the current study. As I discuss in greater detail in Appendix H, the fact that non-response is more prevalent among non-attendees may, based on what we know from the literature, lead us to disproportionately overestimate voter turnout in the non-attendees

[7] i.e. they were aged 25-29 in 2013.

[8] Note also that the ELS sample was recruited not among the general population, but as a representative sample of all US 10th year *high school students* in 2002.

group compared to the college-attendees group (Selb and Munzert 2013; Sciarini and Goldberg 2016; Burden 2000). This mechanism strongly suggests that, in isolation, higher non-response among non-attendees actually biases our estimates in the direction of a negative effective (cf. Appendix H).

F3. Descriptive Statistics

The tables below provide descriptive statistics in the form of univariate and bivariate distributions of the key variables in the main dataset (ELS:2002). Note that statistics are reported for the balanced panel – i.e. for respondents who replied to all waves of the survey.

The dataset includes information on the gender, race/ethnicity, and age of the respondents. Furthermore, the first two waves of data collection also included a parent survey through which it measured parents' education and income. The bottom panel of Table F2 provides descriptive statistics on baseline covariates. In the third and fourth waves of the survey, respondents were asked whether they had moved, whether they still lived with parents, and whether they had gotten married. Table F2 provides descriptive statistics on these time-varying covariates. Both panels also make a comparison between college-goers and non-attendees in terms of these covariates.

Respondents were asked in 2006 whether they voted in the 2004 election, and they were asked in 2012 whether they voted in the 2008 election. As described in the main article, the main independent variable of this study – college attendance – was measured using the administrative postsecondary transcripts. For respondents where transcripts were not available, I supplement with answers from the fourth survey wave. These indicate whether and when the respondent attended postsecondary education. Table F3 shows the mean values of the outcome variable (voting) for the two elections across the groups defined by the main independent variable. Note that the starting points are very different: In 2004, 61% of those who went to college between 2004 and 2008 turned out to vote, whereas only 35% of those who did not go to college voted in 2004. Note also that when we compare the four raw means of Table F3, we see that, descriptively, turnout increased by around 8 percentage points among college-attendees and around 10 percentage points among non-college-attendees, yielding a difference-in-differences of negative 2 percentage points. However, as I argue in the main article and Appendix D, the parallel trends assumption is not tenable when comparing the raw means. This is precisely due to the very different starting points in terms of turnout: The large differences in pre-treatment turnout between attendees and non-attendees would theoretically and methodologically cause

TABLE F2 *Distribution and balance-test on pre-treatment and time-varying covariates by treatment status*

	No College mean	Attended College mean	Difference
Time-varying covariates			
Share that stopped living with parents	0.42	0.60	0.179**
Residential mobility (miles)	51.75	56.48	4.729
Share that moved (residential mobility)	0.32	0.24	-0.088**
Share that got married	0.05	0.02	-0.028**
Pre-treatment covariates			
Share that voted in Pre-treatment Election	0.35	0.61	0.260**
Sex (1=male)	0.52	0.44	-0.082**
Test score math/reading (std. m=50 sd=10)	45.26	53.89	8.634**
Family socio-economic status (std. m=0, sd=1)	-0.38	0.22	0.600**

Note: Data from the balanced panel sample of respondents who replied to all waves of the survey.

a violation of the parallel trends assumption in the raw comparison. Appendix D fully elaborates this argument and presents formal proof that the effect would be biased in the direction of a negative effect if we do not adjust for prior voting. Furthermore, Appendix D3 explicitly

TABLE F3 *Sample means on voting in 2004 and 2008 by treatment status*

	No College (mean)	Attended College (mean)	All Respondents (mean)
Voted in 2004	0.35	0.61	0.58
Voted in 2008	0.45	0.69	0.65
	n = 1187	n = 8200	n = 9387

Note: Data from the balanced panel sample of respondents who replied to all waves of the survey.

compares the raw mean comparison to the main regressions.

Appendix G: On Ethics

This article is based on pre-existing data that was collected by the U.S. Department of Education and that is publicly available. I rely on the public-use data files that have undergone disclosure analyses to protect the confidentiality of the data and minimize disclosure risk. This study only analyzed the data with statistical purposes and took great care to preserve the unidentifiability of the individual respondents. Furthermore, the data collections for both the ELS:2002 and NELS:88 relied on consent from respondents as well as consent from their parents for the high school waves (Ingels et al. 1990, 36; Ingels et al. 2004, 114–116). I therefore consider there to be no ethical issues. Finally, the author affirms that this article adheres to APSA's Principles and Guidance on Human Subject Research.

Appendix H: Overreporting in Self-Reported Turnout

There is a well-established empirical association between educational attainment and overreporting of turnout in surveys compared to validated turnout (Bernstein, Chadha, and Montjoy 2001; Burden 2000; Karp and Brockington 2005; E.R. Hansen and Tyner 2019; Duff et al. 2007). The more educated are more likely to report that they voted even though they did not compared to the less educated. As the current study relies on self-reported turnout from surveys, this deserves thorough attention. More specifically, I consider the following question: (1) What are the consequences of overreporting for the effect estimates found in this study? Afterwards, I adress the concern (2) that the 2008 election constituted an exceptional case where incentives for over-reporting among college-educated were increased. Finally, I provide some evidence to assess the following more descriptive question: (3) What is the extent of overestimation of turnout in the sample used here?

I find that self-reports likely inflate the measured *levels* of turnout in the current study due to non-response and overreporting (Burden 2000; Bernstein et al. 2001), but the individual fixed-effects design cancels out all pre-college differences in overreporting. There may, however, be a causal effect of college on overreporting which could bias this article's estimates upwards. I demonstrate, that in order to be driving the results of the current study, the college effect on overreporting must be larger than the descriptive education differences in overreporting found by prior studies, *and* attenuation effects (e.g. differential non-response increasing turnout inflation among non-college-goers) must be zero. The latter is also implausible based on prior studies.

(1) Potential implications of over-reporting

Firstly, what are the consequences of overreporting for the effect estimates found in this study? As noted above, it is empirically well established that higher education is associated with more overreporting. However, it is important to note that the difference in overreporting between people with longer vs. shorter education can be decomposed into two parts. One part is due to selection into education. In this case, overreporting is higher due to people who take more education simply being different on pre-education characteristics such as political interest and parental participation. The second part is due to the causal effect of education on overreporting. Importantly, the current study is not vulnerable to the former of these two parts: Since we use individual fixed effects (i.e. we study changes in self-reported turnout *within* respondents), we account for the selection portion of overreporting among the highly educated.[9] This means that the extent to which the effect estimates of the current study might be affected by overreporting depends only on the latter part, namely the causal effect of college on overreporting. However, prior studies quite unanimously demonstrate that selection (e.g. based on family characteristics and genetics) accounts for a substantial share of observed associations between education and political variables associated with turnout and overreporting, e.g. political interest and knowledge (Highton 2009; Weinschenk et al. 2021; Weinschenk and Dawes 2019). This leaves us to expect that the consequences of overreporting are not as large in the current study as they would be in studies using other research designs.[10] Furthermore, it suggests that differences in overreporting between educational groups are likely larger than the actual causal effect of education on overreporting (E.R. Hansen and Tyner 2019; Bernstein et al. 2001). In sum, this paper's estimates are not affected by selection-driven over-reporting of turnout. Only a *causal* effect of college on overreporting turnout has the potential to bias the results of the current study towards a positive effect; but three factors determine the existence and severity of this potential bias for the observed results. I assess these factors below:

The first factor is the degree to which the current research design actually accounts for the causal effect of college on overreporting as well. I argue that the current design removes part of the bias that would be due to the causal effect of college since part of the causal effect of college on overreporting was already in place when respondents reported their 2004 voter turnout. Specifically, they report 2004 turnout in the 2006 survey wave of the study. This means that the "treatment group" (college-goers), who attended college between 2004 and 2008, were likely to have already experienced college to some extent, when they, in 2006, reported whether

[9] Formally, since the change in turnout is calculated as the difference between reported 2004 turnout and reported 2008 turnout, the part of overreporting that is due to pre-college factors is absorbed by differencing.

[10] For example, the study by Bernstein et al. (2001) compares the results obtained from multivariate regressions that use reported voting vs. validated voting. They find that the estimated effect of college completion on turnout – compared to having less than 9th grade completed – is 32 percentage points using reported turnout and 27 percentage points using validated turnout. That is a difference in effect size of 5 percentage points. Two reasons suggest that this should be considered an upper bound of the number of percentage points that the effect estimates in the current study would be overestimating due to overreporting. Firstly, the baseline category in the referenced study is less than 9th grade, whereas it is high school completion in the current study. High school's effect on overreporting therefore likely accounts for part of the bias found in Bernstein et al. (2001). Secondly, identification in the multivariate regressions relies solely on adjusting for observed covariates such as race, religion, place of residence, and partisanship. Again, note that prior studies find substantial reduction in estimated education effects when including familial factors, genetic factors, or individual fixed effects. Together, this suggests that yet another part of the 5-percentage-point difference found in Bernstein et al. (2001) is due to selection rather than a causal effect of education.

they voted in 2004 or not. While admittedly relying on respondents' recall is associated with other problems, e.g. reduced measurement reliability, this timing of the survey actually helps us disentangle the causal effect of college on turnout from its effect on overreporting. The fact that both 2004 and 2008 voting were *reported* at a moment in time where some college-goers had attended college means that our difference-in-differences design cancels out part of the effect of college on overreporting.

The second factor that determines the impact of overreporting on the results of this paper is the expected magnitude of the causal effect of college on overreporting. Here, we may rely on prior studies. The study by Bernstein et al. (2001) finds that using validated turnout decreases the effect of college (compared to having less than 9th grade) by 5 percentage points. As mentioned, this is most likely to be an upper bound for the bias in the current study (cf. footnote 9 above) because (1) the estimate likely contains overreporting due to both selection and causation, and (2) the estimate captures both the effect of high school and college. E.R. Hansen and Tyner (2019) use more recent data to study the difference between educational groups in terms of overreporting. In their Figure 2, they compare – among non-voters – the raw difference in overreporting rates between different educational groups. Assuming these differences are the causal effect of education on overreporting, we may translate their estimates among non-voters into an estimate of how much of the increase in reported turnout among college-goers in the current study is actually just overreporting due to the effect of college. These calculations yield point estimates between 1.9 and 6.7 percentage points.[11] Note that the largest difference of 6.7 percentage points is obtained when comparing those who *complete* a college degree to those who complete high school – i.e. when disregarding the "some college" group. This is in contrast to the current study where the "some college" group actually comprises the majority of the college-goers, suggesting that the appropriate estimate should lie in between the two estimates. Furthermore, if we, in the current study, exclude the group of college-completers (where overreporting is expected to be the highest), the effect estimates remain

[11] E.R. Hansen and Tyner (2019) find that *among validated non-voters* and compared to high school completers, overreporting was larger for college-goers by between 5.9 and 17.7 percentage points. Given the reported turnout in the current study, this translates into a growth in fake turnout that is between 1.9 and 6.7 percentage points larger among college-goers than among non-attendees. This is calculated by noting the following relationships:

$$reported\ turnout\ =\ actual\ turnout + overreporting\ rate\ * actual\ non\ voting,$$
$$actual\ turnout = 1 - actual\ non\ voting.$$

Since we know the reported 2008 turnout in the college group (69%), we may make the assumptions outlined above and insert the estimates from E.R. Hansen and Tyner (2019) as "overreporting rate" to solve this set of equations and get the difference between reported and actual turnout.

similar in size and statistical significance. Note also that the argument that the education differences reported in E.R. Hansen and Tyner (2019) are upper bounds also apply here since they are based on raw comparisons of the educational groups. Therefore, the estimates based on E.R. Hansen and Tyner (2019) include the above-mentioned selection portion of educational differences, which the current design already takes into account.

Taken together, when we compare these prior studies' estimated magnitude of overreporting bias to the main causal estimates of the current study, we see that the bias estimates are lower than this study's effect estimates. Specifically, the magnitude of bias found in Bernstein et al. (2001) was 5 percentage points, whereas using E.R. Hansen and Tyner (2019) yielded estimates of 1.9 and 6.7 percentage points. The effect estimates found in the current study are around 9–10 percentage points, with our most conservative estimates being 8.9 percentage points (cf. Table 1 of the main article). In terms of statistical significance, the largest two bias-estimates of 5 and 6.7, if true, would be enough to render the effect estimates insignificant in some of the models.[12] Importantly, however, this is disregarding the ways in which the current design accommodates overreporting bias, as outlined above. Furthermore, differential non-response between the college and non-college groups may counteract some of this bias, as outlined below. Nevertheless, these discussions do support the conclusion that, as mentioned in the main manuscript, further studies should be made that rely on administrative voter data.

Finally, the third factor that determines the severity of overreporting bias for the observed results is the magnitude of counteracting biases that make our estimates more conservative – i.e. mechanisms that pull our estimates in the opposite direction compared to overreporting. Specifically studies suggest that survey non-response is an important such factor. Essentially, two things can be said with confidence non-respondents (i.e. those citizens who do not respond to surveys or drop out of the panel). First, non-respondents are disproportionately non-voters (Selb and Munzert 2013; Sciarini and Goldberg 2016; Burden 2000). That is, our survey consists of disproportionately many (self-reported) voters. Secondly, non-respondents are disproportionately non-college-attendees (cf. Appendix F2 on survey attrition in the current study). If this is true, it means that the non-response causing our survey to contain disproportionately more voters than non-voters is *more* pronounced within the non-college group than within the college-group. This means that we are more likely to overestimate the turnout in 2008 (and thereby the growth in turnout) among non-college-goers than among college-goers. This is because non-college-goers are more likely to drop out between the waves, and those

[12] These are models 4, 5, 6, and 8 of Table 1.

who drop out are disproportionately non-voters. Studies show that overestimation of turnout due non-response is substantial; specific estimates are that it constitutes between half and two thirds of total overestimation (Burden 2000; Sciarini and Goldberg 2016). What are the implications for multivariate studies? Studies in other institutional and demographic contexts than the US suggest that when we try to estimate the effect of education on turnout, the conjunction of bias from overreporting and non-response actually results in more conservative estimates when using surveys than when relying on administrative voting data (Dahlgaard et al. 2019). Thus, both from theoretical arguments and empirical evidence, we would expect that the inflation of turnout due to non-response is substantially larger among the lower education group than among the highly educated group. However, the latter mentioned empirical results are not readily transferred to the context of the current study.[13]

(2) The case of 2008

Not all elections are created equal, and arguably the 2008 election was an extraordinary election in that it saw increased turnout for specific groups defined by e.g. race and age (Philpot, Shaw, and McGowen 2009, see also below). It is plausible that the election of president Barack Obama as the first African American provided a case where incentives for over-reporting among college-educated were increased. If this is the case, the we would expect the increase in self-reported voting to be larger among college-goers than among non-attendees between 2004 and 2008. While this is plausible, there are two important considerations that bolster the current papers conclusions to this concern. First of all, as mentioned above, there is a countervailing mechanism related to the overrepresentation of non-college-goers among those who did not respond to, or dropped out of the survey. To reiterate, as non-responders are more likely to be non-voters, we are over-estimating growth in turnout for the untreated group already. Second, and more importantly, the replication sample of NELS:88 uses the elections in 1992 and 1996 to replicate the findings from the ELS data. Thus, it may serve as a robustness check for whether the increased incentive for over-reporting in 2008 among college-goers is a crucial concern. Reassuringly, in the data from 1992 to 1996 (where Bill Clinton was elected

[13] The current dataset does not allow me to assess whether attending college increased the degree to which respondents feel socially pressured to report voting or socially desirable responding in general. However, increased norms are also one of the hypothesized mechanisms through which college is expected to increase turnout, namely by increasing the social benefits of voting. Moreover, while education may change norms, studies using list experiments have generally found that socially desirable responding only accounts partially or not at all for survey-measured educational differences in attitudes and behaviors more generally (Jeffery 1996; Hofferth 2006; Karp and Brockington 2005). One study even indicates that social desirability bias declines as educational attainment increases (Heerwig and McCabe 2009).

both times), the estimates of the college effect are at least as large as those estimated using self-reported turnout in 2004 and 2008. Nevertheless, the risk of unequal growth in over-reporting between college-goers and non-college-goers remains an important weakness of the current study.

(3) The descriptive extent of over-estimated turnout levels

I close this appendix by providing some descriptive evidence in relation to the following question: (3) What is the extent of overestimation of turnout in the sample used here? Specifically, I make a comparison of estimated turnout levels between studies to assess the extent to which the specific survey used in the current study is particularly problematic or not. There are three caveats when doing this. Firstly, as noted above, the extent to which turnout is overestimated in surveys comprise of both oversampling of voters (non-response) and overreporting by non-voters. Even though we are interested in the latter, we cannot disentangle the two when making the comparisons below. The second caveat is that validated population data on turnout from the US does not enable measuring turnout within groups defined by, e.g., age, race, or education. Since the current study is only nationally representative for a specific age group, we are, therefore, not able to benchmark the observed turnout in the sample against validated population data. We thus have to turn to non-population studies (i.e. samples) to compare the reported turnout rates in the current study against. The third important caveat is the fact that the ELS:2002 data is not sampled to be representative for the general population, but for the cohort of 10th year high school attendees in 2002. Thus, we would expect the mean turnout in the ELS to be biased upward by the exclusion of non-highschool-goers. One benchmark to use is the Current Population Survey's Voting Supplement (CPS), which, for the 2004 presidential election, estimates a turnout rate among 18–29 year-olds of 49.0% when using their regular weights and 45.0% when using the weights suggested in Hur and Achen (2013; McDonald 2020). The raw sample average in the ELS:2002 data used in the current study is a turnout rate of 50.6% in the 2004 election. Turning to the 2008 election, the average turnout among 18-29 year-olds in the CPS was 51.0% with their regular weights. The raw sample average in the ELS:2002 data had increased to 62.3% in 2008. However, these comparisons are arguably too crude because the average age of ELS respondents was 18.5 years in 2004 and 22.4 years in 2008. Thus, first of all, we would expect turnout to increase between 2004 and 2008 simply due to the cohort being older. Second, the CPS average used above uses a fixed age-group. Thus we would expect the gap between ELS and CPS to increase between the two elections, simply because the average age increased in the one source but not in the other. A more rigorous and informative comparison is obtained by comparing the turnout for age groups at the year-level (i.e. 18-year-olds, 19 year-olds and so forth). For this, I draw on age-group specific data from table 1 in Bureau, US Census (2008). The age composition of the ELS data is presented below:

Age in 2004	Share of ELS sample	Age in 2008	Share of ELS sample
18	58%	22	59%
19	36%	23	36%
20	6%	24	5%

In 2004, 58, 36 and 6 percent of the ELS sample were aged 18, 19 and 20 years respectively. Calculating a weighted average of the turnout for these three age groups in CPS yields a rate of 40.3% which we may then compare the ELS rate of 50.6%. In 2008, 59, 36 and 5 percent of the ELS sample were aged 22, 23 and 24 years respectively. Calculating a weighted average of the turnout for these three age groups in CPS yields a rate of 51.1% which we may then compare the ELS rate of 62.3%. This informs us about two things. Firstly, the ELS sample displays higher turnout rates than the corresponding age groups in the CPS data. This may be due to higher over-reporting in the ELS, or due to the fact the the ELS sample is sample of high-school attendees, and not the general population. Secondly, the discrepancy between the CPS and ELS rates, for comparable age groups, is slightly larger in 2008. Specifically the discrepancy is larger by 0.9 percentage points. This change over time in the over-estimation (which is the main threat to identification in the current paper) is reassuringly small.

As mentioned in the main paper, while it seems implausible that differential overreporting is driving the main results, the reliance on self-reported turnout does remain an important limitation to the current study. Consequently, replicating the current panel-based research design using administrative data on voter turnout constitutes a fruitful avenue for future research on the college–participation link.

Appendix I: The Main Article's Analyses: Overview and Details

This appendix provides an overview of each analysis reported in the main article. Appendix I also provides further details on how the analyses were conducted or indicates where in the appendix to find these details. In the order they are described in the article, the analyses are the following (corresponding appendix with further details in parenthesis):

The main analyses and the placebo analysis:

1) Placebo analysis probing the parallel trends assumption (see Appendix A2)
2) Restricted comparison group analyses (see Appendix E)
3) Covariate-by-time fixed-effects analyses (see Appendix I2)
4) Matching analyses (see Appendix I3)
5) Analyses with time-varying covariates (see Appendix D1)

Other analyses:

6) Analyses using the independent replication sample, NELS:88 (see Appendix B)
7) Interaction analyses regarding the compensation hypothesis (see Appendix A1)
8) Power analyses of prior studies and the current study (see Appendix C3)

Below follows a closer elaboration of analyses 3 and 4 listed above.

I2. Covariate-by-Time Fixed Effects

As described in the main article, we may relax the parallel trends assumption by computing counterfactual trends that are specific to groups defined by pre-college covariates (see, e.g., Hall & Yoder 2022; Schafer et al. 2022; Hainmueller & Hangartner 2013). Specifically, the pre-college covariates included are cognitive skills, parents' socio-economic status, gender, and race. In terms of estimation, the group-specific trends are calculated by replacing the time fixed effect of equation (1) with an interaction between the time fixed effect and the pre-college covariates. We then arrive at the following specification (equation 1*), where $\boldsymbol{Z}_i$ is a vector of the pre-college covariates specified above.

$$Y_{it} = \alpha_i + [\beta_t \times \boldsymbol{Z}_i] + \delta[D_i \times PostPeriod_t] + \gamma \boldsymbol{X}_{it} + \epsilon_{it} \qquad (\mathbf{1})*$$

Otherwise, the covariate-by-time fixed effects analyses are estimated using OLS in the same way as the models using simple time fixed effects. Note that a constituent term for $\boldsymbol{Z}_i$ alone would be absorbed in the individual fixed effect α_i. Appendix D2 elaborates on the case where the pre-college covariate is voter turnout. The estimation results and coefficients for all included variables are reported in Appendix A3 along with the other models of Table 1 of the main article.

I3. Matching Analyses

This appendix provides further details on the matching analyses. In Table 1 of the main article, I present results based on a matching difference-in-differences estimator, where I match the treated and untreated groups on pre-treatment characteristics, including sociodemographic factors and prior turnout. The reasoning for including this estimator is that it allows growth in turnout to vary with observed pre-treatment correlates of attending college, such as parents' education (Sant'Anna & Zhao 2020). In other words, matching increases the likelihood that the two groups, prior to treatment, are as similar as possible on characteristics that lead them to change their participation behavior differently over time. In the recent literature on the econometrics of difference-in-differences, the inclusion of a matching estimator has been shown to provide an important kind of robustness, with the basic logic being that by introducing an alternative assumption (getting the matching right), only one of the assumptions (alternative vs. initial parallel trends assumption) has to be fulfilled (Roth et al. 2022; Sant'Anna and Zhao 2020). Below, I elaborate on the choice of matching technique and present pre-matching and post-matching balance statistics. Moreover, I compare the balance and effect estimates from the applied genetic matching to those obtained using propensity scores.

Regarding the matching technique, I rely on genetic matching (Diamond and Sekhon 2013). As is especially clear in the area of education effects, this method has been demonstrated to achieve balance more reliably than, e.g., propensity score matching (Henderson and Chatfield 2011; Mayer 2011). However, as a robustness check, I also run the analysis using inverse probability weighting based on propensity scores, and present estimates and balance obtained this way. The main reason for this is that the difference-in-differences literature cited above is based on this technique. More specifically, since the comparison group is markedly smaller than the treated group, I use genetic matching with replacement (see sample size statistics in Table I3, Panel C below). Due to a smaller effective sample size, this does, however, increase statistical uncertainty, which we observe from the increase in standard errors between the unmatched models 1–3 in Table 1 compared to the matching models. The relative small size of the comparison group and the use of matching with replacement may also increase sensitivity of the results to specification and individual respondents. However, the remarkable similarity across matching and non-matching models in Table 1 of the main article suggests that this is not influencing the results.

Table I3 below reports balance statistics from the genetic matching that underlies the estimates reported in models 7 and 8 of Table 1 in the main article. Specifically, the R package MatchIt was used to perform matching and calculate balance statistics (Ho et al. 2011). The standardized mean difference is calculated based on the standard deviation of the covariate in the treated group. To assess balance, I apply a threshold that the standardized mean difference should be below 0.05. The unmatched data is unbalanced, especially in terms of sex and the two extreme quartiles of SES, where college-goers (treated) have a markedly smaller share of low SES respondents and a larger share of high SES respondents. In terms of race, it is especially the Asian group that is unbalanced, just like pre-college voting is also markedly more prevalent in the college-goer group. Post-matching balance is achieved on all covariates – with perfect balance on all but the Asian and white (non-Hispanic) race categories.

Table I3, Panel A. Balance Statistics: *Unmatched Data*

	Type	Means Treated	Means Control	Std. Mean Diff.
Sex: Male	Binary	0.4384	0.5194	-0.1632
SES 1st Quartile	Binary	0.1592	0.4056	-0.6734
SES 2nd Quartile	Binary	0.2098	0.3314	-0.2985
SES 3rd Quartile	Binary	0.2609	0.1939	0.1525
SES 4th Quartile	Binary	0.3701	0.0691	0.6233
Voted in 2004	Binary	0.3917	0.6509	-0.5311
Race: Hispanic	Binary	0.117	0.1627	-0.1424
Race: Other	Binary	0.1548	0.215	-0.1665
Race: Asian	Binary	0.1023	0.0379	0.2125
Race: White Non-Hispanic	Binary	0.6259	0.5843	0.0859

Table I3, Panel B. Balance Statistics: *Matched* Data (Genetic Matching with Replacement)

	Type	Means Treated	Means Control	Std. Mean Diff.	Diff. vs. Threshold	Std. Pair Diff.
Sex: Male	Binary	0.4384	0.4384	0.000	Balanced, <0.05	0.000
SES 1st Quartile	Binary	0.1592	0.1592	0.000	Balanced, <0.05	0.000
SES 2nd Quartile	Binary	0.2098	0.2098	0.000	Balanced, <0.05	0.000
SES 3rd Quartile	Binary	0.2609	0.2609	0.000	Balanced, <0.05	0.000
SES 4th Quartile	Binary	0.3701	0.3701	0.000	Balanced, <0.05	0.000
Voted in 2004	Binary	0.6083	0.6083	0.000	Balanced, <0.05	0.000
Race: Hispanic	Binary	0.1170	0.1170	0.000	Balanced, <0.05	0.000
Race: Other	Binary	0.1548	0.1548	0.000	Balanced, <0.05	0.000
Race: Asian	Binary	0.1023	0.0709	0.032	Balanced, <0.05	0.104
Race: White Non-Hispanic	Binary	0.6259	0.6574	-0.032	Balanced, <0.05	0.065

Table I3 Panel C. Sample Size Statistics: Genetic Matching with Replacement

Sample Sizes:			
		Control	Treated
All		1186	8198
Matched (Effective Sample Size – ESS)		332.29	8198
Matched		1091	8198
Unmatched		95	0
Discarded		0	0

Robustness Check: Using Propensity Scores

As mentioned above, I compare the effect estimates and balance that I obtain using the genetic matching method to those obtained from using inverse probability weighting based on propensity scores because this method is prevalent in the cited difference-in-differences literature (Sant'Anna and Zhao 2020). Table I4, Panel A below reports balance statistics using this method. The R package WeightIt was used to calculate propensity scores, weights, and balance statistics (Greifer 2022). Firstly, this method also achieves balance on all covariates based on the 0.05 threshold for standardized mean differences. In terms of differences compared to the genetic matching, we may observe that no covariates achieved a difference of 0 using propensity scores, whereas this was the case for 8 out of 10 covariates using genetic matching. More specifically, the "Hispanic" and "Other" race categories as well as the two upper quartiles of socioeconomic status have a standardized mean difference above .018 in the propensity score technique but 0 using genetic matching. The effective sample size is similar using the propensity score method (cf. Table I4, Panel B).

Finally, Table I5 reports the effect estimates obtained when using the two different kinds of matching techniques. The results are remarkably similar across methods.

Table I4, Panel A. Balance Statistics: *Matched* Data (Inverse Probability Weighting with Propensity Scores)

	Type	Std. Mean Diff.	Diff. vs. Threshold
Sex: Male	Binary	-0.0085	Balanced, <0.05
SES 1st Quartile	Binary	-0.0021	Balanced, <0.05
SES 2nd Quartile	Binary	0.0094	Balanced, <0.05
SES 3rd Quartile	Binary	-0.0257	Balanced, <0.05
SES 4th Quartile	Binary	0.0185	Balanced, <0.05
Voted in 2004	Binary	0.0053	Balanced, <0.05
Race: Hispanic	Binary	-0.0231	Balanced, <0.05
Race: Other	Binary	-0.0247	Balanced, <0.05
Race: Asian	Binary	0.0070	Balanced, <0.05
Race: White Non-Hispanic	Binary	0.0408	Balanced, <0.05

Table I4, Panel B. Sample Size: Inverse Probability Weighting Based on Propensity Scores

Effective Sample Sizes (ESS):			
		Control	Treated
Unadjusted ESS		1186	8198
Adjusted ESS		371.58	8198
Discarded		0	0

TABLE I5 *Robustness of Matching Analyses using Genetic Matching or Propensity Scores*

	Matched (Genetic)		Matched (Propensity Score)	
	(1)	(2)	(3)	(4)
No College	ref.	ref.	ref.	ref.
Attended College	0.116***	0.101***	0.107***	0.109***
	(0.026)	(0.027)	(0.025)	(0.025)
Time FEs & Individual FEs	✓	✓	✓	✓
Pre-college Voting × Time FEs	✓	✓	✓	✓
Time-varying Controls		✓		✓
Units	9,289	8,582	9,384	8,751
Observations	18,578	17,068	18,768	17,397

Note: Difference-in-differences estimates of how college-goers changed their turnout compared to respondents who did not go to college between the 2004 and 2008 elections. The models compare college attendees to a matched comparison group. Models 1 and 2 uses Genetic Matching with replacement, as reported in Table 1 of the main article. Models 3 and 4 use Inverse Probability Weighting based on Propensity Scores. Matching covariates are pre-college race, gender, prior voting and parental SES). Time-varying controls include residential mobility, living with parents, getting married, crime victimization, becoming seriously ill or disabled, labor market status, becoming a parent, parental divorce, parental job loss, serious illness in the family, parent died and relative or friend died. Robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Appendix J: Prior Studies on the Compensation Hypothesis and Temporal Heterogeneity in the Effect of College on Voter Turnout

In this appendix, I elaborate on the theoretical and empirical arguments underlying the so-called compensation hypothesis that I investigate in the results section of the main article. As outlined in the article, college education is expected to impact political participation mainly by providing students with resources that are relevant for participation (Brady, Verba, and Schlozman 1995). In the short run, these include relevant knowledge, organizational skills, and social relations that enable citizens to follow through on their intention to vote (Hill 2018; Holbein 2017). These resources may arguably exhibit ceiling effects, such that those who have already obtained politically relevant knowledge and norms of civic duty through political socialization in the family will not benefit as much from obtaining these resources in college. Conversely, political participation would increase for those who do not have prior experience with politics. This is the so-called compensation hypothesis that education effects on political participation are mainly constituted by increasing the level of participation among less privileged students (Neundorf, Niemi, and Smets 2016; Lindgren, Oskarsson, and Persson 2019). Since the share of US youth who attend college has almost doubled in the last 40 years, the student body has broadened and diversified.

With the compensation hypothesis in mind, this is a situation that actually enables college education to play a compensating role. Taken together, this suggests that the effect of college on college-goers may have changed over the decades, where effects would be larger in decades with a more diverse population of college-goers. In fact, there is tangible empirical evidence to bear on the notion that the effect of college may have changed over time. Burden (2009) shows how the association between college education and voter turnout became stronger in the 1980s. Furthermore, Burden (2009) suggest that this may be due to supply-side reasons, i.e. that the content of education and the skills it provides changed over time. The reason suggested here, which may supplement the arguments of Burden (2009), relates to the student population: College-goers in earlier decades are more likely to represent a more priviledged stratum of society (Stevens et al. 2008). Importantly, this is a group that theoretically, due to the compensation hypothesis, may be less prone to the civic effects of education.

The findings of the current paper suggest that future research may fruitfully investigate the extent of this temporal heterogeneity and the reasons that underlie it. Based on my results, it may be that the compensating role of education also changes over time. It may also be that we have left to discover what it is that higher education compensates for. As mentioned in the main article, I discuss this in Appendix B.

To reiterate, the argument that the effect of college on college-goers would increase (over time) as college expands to less privileged groups is a combination of two well established arguments, namely 1) the compensation hypothesis and 2) the increasing representation of less privileged students among college-goers. Individually, these arguments are featured in prior work (e.g. Neundorf, Niemi, and Smets 2016; Trow 2007; Burden 2009; Stevens, Armstrong, and Arum 2008). I argue that if both of the two arguments hold, we would expect the average effect of college on college-goers to increase over time and this may explain the mixed pattern of findings in the literature. In an important study, Lindgren, Oskarsson, and Persson (2019) also highlight these two component arguments in conjunction (which they label "the return effect" and "the resource effect", respectively) as two ways in which rising overall education levels (due to educational reforms) may reduce class inequality in political participation. The contribution of the current paper is to combine the two arguments to derive their implications for how the effect of college will change over time.

Appendix K: The Use of OLS/Linear Probability Model in the Current Study and in Panel Studies with Binary Outcomes in General

The current study follows the standard among recent studies using panel data with binary outcomes in using a linear probability model (Sønderskov et al. 2020; Hall and Yoder 2021; Hainmueller and Hangartner 2013). More specifically, as outlined in the "Analytical Strategy and Estimation" section of the main article, the current study relies on estimating the two-way fixed-effects model of equation (1) using ordinary least squares (OLS). Below, I outline the two main reasons for choosing a linear model in the current case – i.e. specifically when having panel data with two time periods and a binary outcome variable. These arguments also suggest that, e.g., a logit model for limited dependent variables should not be applied in this particular case.

First of all, in the current case where treatment, outcome, and time-periods are binary, the linear probability model is simply a comparison of conditional means on the turnout variable. Thereby, the linear probability model does not impose any parametric constraints on the functional form of the relationship between the treatment and the outcome. Therefore, the standard reasons for preferring, e.g., a logit model to OLS do not apply in this case (Stock and Watson 2015, 386).

Second and more importantly, when we have panel data and binary outcomes, logit estimation in the context of panel data drops all respondents that have the same outcome in both time periods. This makes the linear probability model preferable over a binary model (Sønderskov et al. 2020). Not only does a binary model remove a large part of respondents, but this also constitutes post-treatment conditioning (Elwert and Winship 2014).

Appendix L: Effects for Non-Presidential Elections

As outlined in the main article, the timing of the ELS:2002 cohort sample is particularly suited for providing a well-powered estimation of the effect of college on presidential turnout for those who attended college between 2004 and 2008. However, the ELS survey also contains self-reports of voting in local elections. Specifically, in 2006 respondents were asked if they had voted in a local election within the previous two years. In 2012, respondents were asked if they had voted in a local election within the previous three years. Thus, to estimate the effect of attending college on participation in local elections, we may look at the subsample of respondents who did not attend college at all before 2006, but either attended college between 2006 and 2009 (treated) or did not (untreated). Making this comparison reduces the sample size significantly, because many respon dents enrolled between 2004 and 2006. This also means that the estimates of the college effect on college-goers (ATT) obtained in relation to local elections are local to a specific group of respondents, namely those who started college later than most college-goers. Thus while this serves as a test of the theoretical notion that college increases local election turnout, we should be cautious to compare the substantial size of the coefficient to that obtained in the main analyses in table 1 of the main paper. Table L1 presents the difference-in-difference estimates of this analysis. I use the same specification as equation (1) but with the treatment and outcome variables changed as outliend above. In both specifications I find a positive and statistically significant effect of attending college on the probability of reporting to vote in a local election of around 5.5 percentage points. Given the inhibiting factors for comparisons with the papers main analyses, this analyses mostly speaks to the generalizability of a college effect: I find that college increases participation for college-goers both in local and presidential elections in the short term.

TABLE L1 *Effect of College Education on Non-presidential Turnout. Sample: Those who did not attend college before 2006*

	(1)	(2)
No College	ref.	ref.
Attended College	0.057*	0.054*
	(0.025)	(0.026)
Time FEs & Individual FEs	✓	✓
Pre-college Voting × Time FEs	✓	✓
Additional Pre-college Covariates × Time FEs		✓
Units	1,558	1,514
Observations	2,894	2,813

Note:Difference-in-differences estimates of how college-goers changed their turnout in local elections (non-presidential) elections compared to respondents who did not go to college between the 2006 and 2009 elections. All models include time trends that are specific to whether a respondent voted in the pre-college period or not. Model 2 further interacts time fixed effects with pre-college cognitive skills, gender, race, and parental education and income. OLS estimates with robust standard errors clustered by individual in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Appendix References

Apfeld, Brendan, Emanuel Coman, John Gerring, and Stephen Jessee. 2021. "Education and Social Capital." *Journal of Experimental Political Science*: 1-27.

Berinsky, Adam J., and Gabriel S. Lenz. 2011. "Education and Political Participation: Exploring the Causal Link." *Political Behavior* 33(3): 357-373.

Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. "Overreporting Voting." *Public Opinion Quarterly* 65(1): 22-44.

Bloom, Howard S. 1995. "Minimum Detectable Effects." *Evaluation Review* 19(5): 547-556.

Brady, Henry E., Sidney Verba, and Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *American Political Science Review* 89(2): 271-294.

Bureau, US Census (2008). Voting and Registration in the Election of November 2008. Census.Gov., Retrieved from https://www.census.gov/library/publications/2012/demo/p20-562rv.html

Burden, Barry C. 2000. "Voter Turnout and the National Election Studies." *Political Analysis* 8(4): 389-398.

---. 2009. "The dynamic effects of education on voter turnout." *Electoral Studies* 28(4): 540-549.

Burden, Barry C., Jason M. Fletcher, Pamela Herd, Bradley M. Jones, and Donald P. Moynihan. 2017. "How Different Forms of Health Matter to Political Participation." *The Journal of Politics* 79(1): 166-178.

Burden, Barry C., Pamela Herd, Bradley M. Jones, and Donald P. Moynihan. 2020. "Education, early life, and political participation: New evidence from a sibling model." *Research & Politics* 7(3): 205316802095831.

Christopher, Elise M. 2014. *ELS:2002 Postsecondary Education Transcript Study: Addendum to the Public-Use File (PETS Public-use Read-me)*. Washington, DC.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Coppock, Alexander, and Donald P. Green. 2016. "Is Voting Habit Forming?" *American Journal of Political Science* 60(4): 1044-1062.

Curtin, Thomas R., Steven J. Ingels, Shiying Wu, Ruth Heuer, and Jeffrey Owings. 2002. *National Education Longitudinal Study of 1988: Base-Year to Fourth Follow-up Data File User's Manual* Washington, DC.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Dahlgaard, Jens Olav, Jonas Hedegaard Hansen, Kasper M. Hansen, and Yosef Bhatti. 2019. "Bias in Self-reported Voting and How it Distorts Turnout Models." *Political Analysis* 27(4): 590-598.

Dahlgaard, Jens Olav, and Kasper M. Hansen. 2021. "Twice the Trouble: Twinning and the Cost of Voting." *The Journal of Politics* 83(3): 1173-1177.

Debell, Matthew, Jon A. Krosnick, Katie Gera, David S. Yeager, and Michael P. McDonald. 2018. "The Turnout Gap in Surveys." *Sociological Methods & Research*: 004912411876908.

Dee, Thomas S. 2004. "Are there civic returns to education?" *Journal of Public Economics* 88(9-10): 1697-1720.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188.

Diamond, Alexis, and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95(3): 932-945.

Dinesen, Peter Thisted, Christopher T. Dawes, Magnus Johannesson, Robert Klemmensen, Patrik Magnusson, Asbjørn Sonne Nørgaard, Inge Petersen, and Sven Oskarsson. 2016. "Estimating the Impact of Education on Political Participation: Evidence from Monozygotic Twins in the United States, Denmark and Sweden." *Political Behavior* 38(3): 579-601.

Duff, B., M. J. Hanmer, W. H. Park, and I. K. White. 2007. "Good Excuses: Understanding Who Votes With An Improved Turnout Question." *Public Opinion Quarterly* 71(1): 67-90.

Elwert, Felix, and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology* 40(1): 31-53.

Engelman, Michal, Won-Tak Joo, Jason Fletcher, and Barry Burden. 2021. "Health, Wealth, and Voting Trajectories in Later Life." *The Journals of Gerontology: Series B*. 77(4): 827-837.

Gerber, Alan S., Donald P. Green, and Ron Shachar. 2003. "Voting May Be Habit-Forming." *American Journal of Political Science* 47(3): 540-550.

Gidengil, Elisabeth, Lasse Tarkiainen, Hanna Wass, and Pekka Martikainen. 2019. "Turnout and Education: Is Education Proxying for Pre-Adult Experiences Within the Family?" *Political Science Research and Methods* 7(2): 349-365.

Greifer, Noah. 2022. *"WeightIt: Weighting for Covariate Balance in Observational Studies."* Accessed 2022-01-25. https://ngreifer.github.io/WeightIt/.

Hainmueller, Jens, and Dominik Hangartner. 2013. "Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination." *American Political Science Review* 107(1): 159-187.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27(2): 163-192.

Hall, Andrew, and Jesse Yoder. 2021. "Does Homeownership Influence Political Behavior? Evidence from Administrative Data." *The Journal of Politics*.

Hansen, Eric R., and Andrew Tyner. 2019. "Educational Attainment and Social Norms of Voting." *Political Behavior*.

Hansen, J.H. 2016. "Residential Mobility and Turnout: The Relevance of Social Costs, Timing and Education." *Political Behavior* 38(4): 769-791.

Heerwig, Jennifer A., and Brian J. McCabe. 2009. "Education and Social Desirability Bias: The Case of a Black Presidential Candidate." *Social Science Quarterly* 90(3): 674-686.

Henderson, John, and Sara Chatfield. 2011. "Who Matches? Propensity Scores and Bias in the Causal Effects of Education on Participation." *The Journal of Politics* 73(3): 646-658.

Highton, Benjamin. 2009. "Revisiting the Relationship between Educational Attainment and Political Sophistication." *The Journal of Politics* 71(4): 1564-1576.

Highton, Benjamin, and Raymond E. Wolfinger. 2001. "The First Seven Years of the Political Life Cycle." *American Journal of Political Science* 45(1): 202.

Hill, Seth J. 2018. "Following Through on an Intention to Vote: Present Bias and Turnout." *Political Science Research and Methods*: 1-8.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2011. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." *Journal of Statistical Software* 42(8): 1 - 28.

Hofferth, Sandra L. 2006. "Response Bias in a Popular Indicator of Reading to Children." *Sociological Methodology* 36(301-315.

Holbein, John B. 2017. "Childhood Skill Development and Adult Political Participation." *American Political Science Review* 111(3): 572-583.

Holbein, J. B., Rangel, M. A., Moore, R., & Croft, M. (2023). Is Voting Transformative? Expanding and Meta-Analyzing the Evidence. *Political Behavior*, 45(3), 1015–1044.

Hur, Aram, and Christopher H. Achen. 2013. "Coding Voter Turnout Responses in the Current Population Survey." *Public Opinion Quarterly* 77(4): 985-993.

Ingels, Steven J., Sameer Y. Abraham, Rosemary Karr, Bruce D. Spencer, Martin. R Franke, and Jeffrey Owings. 1990. *"National Education Longitudinal Study of 1988 Base Year: Student Component Data File User's Manual."* National Center for Education Statistics (NCES). Accessed 2022-02-01. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=90464.

Ingels, Steven J., Daniel J. Pratt, Christopher P. Alexander, Donna M. Jewell, Erich L. Lauff, Tiffany L. Mattox, and David Wilson. 2014. *Education Longitudinal Study of 2002 Third Follow-up Data File Documentation (NCES 2014-364)*. Washington, DC.: National Center for Education Statistics, U.S. Department of Education.

Ingels, Steven J., Daniel J. Pratt, James. E. Rogers, Peter H. Siegel, Ellen S. Stutts, and Jeffrey Owings. 2004. *"Education Longitudinal Study of 2002: Base Year Data File User's Manual "*. National Center for Education Statistics (NCES). Accessed 2022-02-01. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2004405.

Jeffery, R. W. 1996. "Bias in reported body weight as a function of education, occupation, health and weight concern." *Addict Behav* 21(2): 217-22.

Kam, Cindy D., and Carl L. Palmer. 2008. "Reconsidering the Effects of Education on Political Participation." *The Journal of Politics* 70(3): 612-631.

Karp, Jeffrey A., and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *The Journal of Politics* 67(3): 825-840.

Kiley, Kevin, and Stephen Vaisey. 2020. "Measuring Stability and Change in Personal Culture Using Panel Data." *American Sociological Review* 85(3): 477-506.

Lindgren, Karl-Oskar, Sven Oskarsson, and Mikael Persson. 2019. "Enhancing Electoral Equality: Can Education Compensate for Family Background Differences in Voting Participation?" *American Political Science Review* 113(1): 108-122.

Mayer, Alexander K. 2011. "Does Education Increase Political Participation?" *The Journal of Politics* 73(3): 633-645.

McDonald, Michael P. 2020. "Voter Turnout Demographics - United States Elections Project."

NCES. 2002. *"NELS:88/2000 Public Use Data Files and Electronic Codebook - Base Year through Fourth Follow-up."* [Computer file]. National Center for Education Statistics (NCES). Accessed 2022-02-01. https://nces.ed.gov/onlinecodebook.

---. 2015. *"Education Longitudinal Study of 2002 (ELS:2002) Postsecondary Transcripts Public-use Data File."* [Computer file]. National Center for Education Statistics (NCES). Accessed 2022-02-01. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2015314.

Neundorf, Anja, Richard G. Niemi, and Kaat Smets. 2016. "The Compensation Effect of Civic Education on Political Engagement." *Political Behavior* 38(4): 921-949.

Persson, Mikael. 2014. "Testing the Relationship Between Education and Political Participation Using the 1970 British Cohort Study." *Political Behavior* 36(4): 877-897.

---. 2015. "Education and Political Participation." *British Journal of Political Science* 45(3): 689-703.

Philpot, Tasha S., Daron R. Shaw, and Ernest B. McGowen. (2009) "Winning the race: Black voter turnout in the 2008 presidential election." *Public Opinion Quarterly* 73(5): 995-1022.

Plutzer, Eric. 2002. "Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood." *American Political Science Review* 96(1): 41-56.

Rambachan, A., & Roth, J. (2023). A More Credible Approach to Parallel Trends. *Review of Economic Studies*, 90(5), 2555–2591.

Sant'Anna, Pedro H. C., and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of Econometrics* 219(1): 101-122.

Sciarini, Pascal, and Andreas C. Goldberg. 2016. "Turnout Bias in Postelection Surveys: Political Involvement, Survey Participation, and Vote Overreporting." *Journal of Survey Statistics and Methodology* 4(1): 110-137.

Sears, David O, and Carolyn L Funk. 1999. "Evidence of the long-term persistence of adults' political predispositions." *The Journal of Politics* 61(1): 1-28.

Selb, Peter, and Simon Munzert. 2013. "Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys." *Electoral Studies* 32(1): 186-196.

Slough, T., & Tyson, S. A. (2023). External Validity and Meta-Analysis. *American Journal of Political Science*, 67(2), 440–455.

Sønderskov, Kim Mannemar, Peter Thisted Dinesen, Steven E. Finkel, and Kasper M. Hansen. 2020. "Crime Victimization Increases Turnout: Evidence from Individual-Level Administrative Panel Data." *British Journal of Political Science*: 1-9.

Stevens, Mitchell L., Elizabeth A. Armstrong, and Richard Arum. 2008. "Sieve, Incubator, Temple, Hub: Empirical and Theoretical Advances in the Sociology of Higher Education." *Annual Review of Sociology* 34(1): 127-151.

Stock, James H., and Mark W. Watson. 2015. *Introduction to econometrics*. Updated third edition, Global edition ed. Vol. Book, Whole*The Pearson series in economics; Always learning*. Boston: Pearson.

Stubager, Rune. 2008. "Education effects on authoritarian–libertarian values: a question of socialization." *The British Journal of Sociology* 59(2): 327-350.

Tenn, Steven. 2005. "An Alternative Measure of Relative Education to Explain Voter Turnout." *The Journal of Politics* 67(1): 271-282.

---. 2007. "The Effect of Education on Voter Turnout." *Political Analysis* 15(4): 446-464.

Weinschenk, Aaron C., and Christopher T. Dawes. 2019. "The Effect of Education on Political Knowledge: Evidence From Monozygotic Twins." *American Politics Research* 47(3): 530-548.

Weinschenk, Aaron C., Christopher T. Dawes, Stig Hebbelstrup Rye Rasmussen, and Robert Klemmensen. 2021. "The relationship between education and political knowledge: evidence from discordant Danish twins." *Journal of Elections, Public Opinion and Parties*: 1-13.