APPENDIX ONLINE

Liesbet Hooghe, Gary Marks, Jonne Kamphorst. 2024. "Why it makes sense to pay attention to a person's field alongside their level of education: Voting on the socio-cultural divide." *American Political Science Review*.

PLEASE NOTE: FURTHER DOCUMENTATION AND DATA FILES CAN BE FOUND IN THE APSR DATAVERSE at https://doi.org/10.7910/DVN/YMLR66

Contents

A. Descriptive support for the ESS analysis

Table A.1. Country coverage and sample size

Country	Wave 2 (2004)	Wave 3 (2006)	Wave 4 (2008)
Austria	1,044	1,249	no wave
Belgium	1,177	1,200	1,182
Denmark	1,115	1,196	1,265
Finland	1,268	1,215	1,368
Germany	1,649	1,714	1,734
Greece	1,292	no wave	1,232
France	935	979	1,099
Ireland	1,480	992	1,156
Netherlands	1,380	1,393	1,384
Norway	1,281	1,257	1,111
Portugal	917	1,112	1.034
Spain	926	998	1,376
Sweden	1,401	1,414	1,339
Switzerland	926	745	752
United Kingdom	1,073	1,461	1,372
Total	17,864	16,925	17,404

Note: The sample encompasses country-waves in which vote choice includes a GAL or TAN political party. Light-gray shaded cells are country-waves where both a GAL and a TAN party were presented as options in a particular ESS wave; medium-gray cells indicate where only GAL competed; there was no country-wave with a TAN party but no GAL party. Wave 3 was not fielded in Greece, and wave 4 was not fielded in Austria. N=52,193. Detailed information on our coding of GAL and TAN parties is available on Dataverse (Additional Documentation).

Table A.2a. Field of education: CECT scores by field of study

This table lists the average CECT, which reflects the relative preponderance of human-centered education, for each of fourteen fields of study offered to respondents in the European Social Survey. The scores are rescaled 0-1 for ease of interpretation. Each respondent receives the score that matches their declared field of study of their highest completed degree. Please refer to the special do file "CECT_ESS.do" online that contains separate code for integrating these CECT scores into the ESS dataset.

Educational field	**Educational CECT**	**Size of the field**
Teacher training, education	1.000	7.04
Arts, fine/applied	0.952	2.3
Humanities	0.952	4.1
Social studies/administration/media/culture	0.861	5.3
Personal care services	0.680	5.8
Science/mathematics/computing etc.	0.614	4.8
Medical/health services/ nursing etc.	0.554	10.6
General education	0.531	17.8
Public order and safety	0.494	1.1
Law and legal services	0.312	1.5
Economics/commerce/business administration	0.188	15.3
Technical and engineering	0.036	19.7
Transport and telecommunications	0.036	1.5
Agriculture/forestry	0.000	3.1
Mean / Total	**0.441**	**100**

Note: N=35198 respondents in the sample, who indicated a field of study. Respondents with only primary education or no degree receive a score of 0.361.

Table A.2b. Key independent variables

Educational CECT	Human-centered content of a person's field of education. CECT is the ratio of communicative plus cultural skills to the sum of four skill categories for a given field of education (rescaled 0 to 1). Skill ratios are derived from van de Werfhorst & Kraajkamp (2001: 301) and allocated to each of 14 fields of study in the ESS (*edufld*): 1) general/no specific field; 2) arts, fine/applied; 3) humanities; 4) technical and engineering; 5) agriculture/forestry; 6) teacher training/education; 7) science/mathematics/computing etc.; 8) medical/health services/nursing etc.; 9) economics/commerce/business administration; 10) social studies/administration/ media/culture; 11) law and legal services; 12) personal care services; 13) public order and safety; 14) transport and communications.

Occupational CECT	Average CECT of all respondents in a person's ISCO-3 level occupation (using ESS variable *iscoco* and our Educational CECT variable) (rescaled 0 to 1). We apply an occupational CECT score to each individual in that occupation.
Field income	Average income of all respondents in a person's field by level of education (rescaled 0 to 1). We are using ESS variable *hinctnt* for round 2-3 and *hinctnta* for round 4 for income; ESS variable *edufld* for field of education, and ESS variable *edulvla* for level of education.
Level of education	A five-category variable (ESS variable *edulvla*), whereby 1=less than lower secondary education (ISCED 0-1); 2 = lower secondary education completed (ISCED 2); 3 = upper secondary education completed (ISCED 3); 4 = post-secondary non-tertiary education completed (ISCED 4); 5 = tertiary education completed (ISCED 5-6).
Higher education	A dichotomous variable that takes on a value of 1 if the respondent completed post-secondary non-tertiary education (categories 4 and 5 of *edulvla*) and zero otherwise.

Table A.2c. Controls

Female	Self-reported (ESS variable *gndr*), recoded to male=0, and female=1.
Income	Ten-category variable that summarizes income in deciles. ESS introduced deciles in wave 4. We use ESS documentation on country-specific calculation of the deciles to transform income information from 2004 and 2006 to the new decile measure (ESS variable *hinctnt* for round 2-3, *hinctnta* for round 4).
Rural	Five-category variable that reports respondent's self-description of the area where they live (ESS variable *domicil*), ranging from 1=big city; 2=suburbs or outskirts of big city; 3=town or small city; 4=country village; 5=farm or home in countryside.
Secular	Seven-category variable tapping attendance of religious services (ESS variable *rlgatnd*), ranging from 1=every day; 2=more than once a week; 3=once a week; 4=at least once a month; 5=only on special holy days; 6=less often; 7=never.
Age	Calculation bases on year of birth (ESS variable *agea*).
Occupation	Classifies a person's job or past job in eight categories using information on employment relationship, work logic, and job content derived from ISCO-88 and estimated by applying the coding scheme developed by Oesch (2006): 1) semi-employed professionals and large employers; 2) small business owners; 3) technical (semi-)professionals; 4) production workers; 5) (Associate) managers; 6) clerks; 7) socio-cultural (semi-)professionals; 8) service workers.
Country	AT, BE, CH, DE, DK, FI, FR, GR, IE, NL, NO, PO, ES, SV, UK
Time	ESS waves 2002, 2004, 2006

Note: Descriptives and further information are available on Dataverse (Additional Documentation).

B. The baseline model of field of education

The baseline model shows that both field variables are highly significant under controls (Figure 2).

Table A.3. Baseline model

	GAL	TAN
Educational CECT	0.780***	-0.372***
	(0.074)	(0.097)
Occupational CECT	0.900***	-0.972***
	(0.125)	(0.196)
EDUCATIONAL CONTROLS		
Field income	-0.089	-0.320
	(0.305)	(0.683)
Level of education		
Did not complete lower secondary degree	*Reference*	*Reference*
Lower secondary degree	0.111	0.165
	(0.144)	(0.163)
Higher secondary degree	0.444**	-0.015
	(0.175)	(0.275)
Post-secondary degree	0.795***	-0.219
	(0.211)	(0.275)
Tertiary degree	0.912***	-0.879***
	(0.235)	0.294
OTHER CONTROLS		
Female	0.123***	-0.251***
	(0.039)	(0.058)
Income	-0.048***	-0.047***
	(0.008)	(0.012)
Rural	-0.178***	0.019
	(0.015)	(0.021)
Age	-0.021***	-0.009***
	(0.001)	(0.002)
Secular	0.215***	0.138***
	(0.015)	(0.020)
Country intercept variance	0.448***	1.732**
	(0.168)	(0.761)
ISCO intercept variance	0.095***	0.116***
	(0.020)	(0.029)
Intercept	-3.161***	-2.097***
	(0.249)	(0.585)
Observations	40,943	30,794
Groups	15	11
Log Likelihood	-11672.8	-6515,9
BIC	23526.2	13207.5

Note: The coefficients are log odds derived from a multilevel mixed-effects logistic models with oim clustering by country and ISCO-3 categories, with time fixed effects. Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

C. Testing a field logic of occupation

This section provides background models for assessing the relative strength of an occupational logic and the proposed field logic of occupation. Occupations are operationalized by means of the Oesch (2006) categorization. The field logic is operationalized by imputing for each respondent the average CECT of all respondents in their occupation; occupations are aggregated at the ISCO-3 level. The models in Table A.4 produce Figure 3 in the article. The reference category is production workers.

Table A.4. Field of education and occupation

	GAL	TAN
EDUCATIONAL FIELD LOGIC		
Educational CECT	0.781***	-0.361**
	(0.075)	(0.097)
Occupational CECT	0.746***	-0.800***
	(0.183)	(0.271)
OCCUPATIONAL LOGIC		
Production workers	*Ref. category*	*Ref. category*
Self-employed profs & large employers	0.070	-0.379*
	(0.123)	(0.197)
Small business owners	-0.040	-0.192**
	(0.089)	(0.092)
Technical (semi-)professionals	0.253***	-0.573***
	(0.090)	(0.119)
(Associate) managers	-0.020	-0.571***
	(0.081)	(0.103)
Clerks	-0.144	-0.194*
	(0.091)	(0.109)
Socio-cultural (semi-)professionals	0.084	-0.478***
	(0.106)	(0.168)
Service workers	-0.057	0.087
	(0.083)	(0.099)
CONTROLS		
Field income	-0.140	0.201
	(0.309)	(0.685)
Level of education		
Did not complete lower secondary education	*Ref. category*	*Ref. category*
Lower secondary degree	0.130	0.102
	(0.145)	(0.162)
Higher secondary degree	0.469***	-0.031
	(0.176)	(0.201)
Post-secondary degree	0.715***	-0.259
	(0.212)	(0.276)
Tertiary degree	0.906***	-0.842***
	(0.237)	(0.294)
Female	0.142***	-0.270***
	(0.040)	(0.059)
Income	-0.049***	-0.036***

	(0.009)	(0.012)
Rural	-0.178***	0.015
	(0.015)	(0.021)
Age	-0.020***	-0.008***
	(0.001)	(0.002)
Secular	0.216***	0.140***
	(0.015)	(0.020)
Country intercept variance	0.447***	1.733**
	(0.168)	(0.761)
ISCO intercept variance	0.085***	0.070***
	(0.019)	(0.023)
Intercept	-3.108***	-2.463***
	(0.254)	(0.588)
Observations	40,762	30,654
Groups	15	11
Log likelihood	-11622.0	-6459.0
BIC	23498.7	13165.9

Note: The coefficients are log odds, derived from a multilevel mixed-effects logistic models with OIM clustering by country and ISCO-3 categories, with time fixed effects not shown. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

C. Testing the effect of field among higher and lower educated

Table A.5 interacts field of education with level of education to produce Figure 4 in the article. For both higher and lower educated individuals, field is a significant predictor of vote choice.

Table A.5. Field of education for higher and lower educated

	GAL	TAN
Educational CECT	0.665***	-0.246**
	(0.098)	(0.109)
Occupational CECT	0.868***	-0.933***
	(0.127)	(0.200)
Higher education	0.152*	-0.547***
	(0.087)	(0.109)
Higher education x educational CECT	0.325***	-0.499***
	(0.121)	(0.180)
CONTROLS		
Field income	0.751***	-0.763**
	(0.163)	(0.365)
Female	0.129***	-0.263***
	(0.039)	(0.058)
Income	-0.047***	-0.049***
	(0.008)	(0.012)
Rural	-0.179***	0.022
	(0.015)	(0.021)
Age	-0.021***	-0.009***
	(0.001)	(0.002)
Secular	0.215***	0.138***

	(0.015)	(0.020)
Country intercept variance	0.453***	1.780**
	(0.170)	(0.782)
ISCO intercept variance	0.098***	0.125***
	(0.020)	(0.030)
Intercept	-3.208***	-1.710***
	(0.248)	(0.535)
Observations	40,943	30,794
Groups	15	11
Log Likelihood	-11678.0	-6525.9
BIC	23515.4	13206.9

Note: The coefficients are log odds, derived from a multilevel mixed-effects logistic models with oim clustering by country and ISCO-3 occupational categories, time fixed effects not shown. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

D. Testing how gender and field of education relate

Table A.6 provides the background models for Figure 4 in the article. Models 1 and 4 compare models that include field but exclude gender against the baseline model that has both (models 2 and 5); this shows that the effect of field is strongly resilient to including gender. Models 3 and 6 include gender but exclude field; this reveals that the gender gap is overestimated if information on field of education is omitted. A model that includes both variables (models 2 and 5) is superior to one that includes gender only (models 3 and 6), as indicated by the lower BIC for the former.

Table A.6. The effect of field -- with or without controls for gender

	Voting GAL			Voting TAN		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Educational CECT	0.812***	0.780***		-0.453***	-0.372***	
	(0.074)	(0.074)		(0.096)	(0.097)	
Occupational CECT	0.972***	0.900***		-1.204***	-0.972***	
	(0.123)	(0.125)		(0.190)	(0.196)	
Female		0.123***	0.298***		-0.251***	-0.426***
		(0.039)	(0.038)		(0.058)	(0.054)
CONTROLS						
Field income	-0.105	-0.089	-1.855***	-0.359	-0.320	0.393
	(0.304)	(0.305)	(0.280)	(0.684)	(0.683)	(0.665)
Level of education						
No lower secondary degree	*Ref. category*	*Ref. category*	*Ref. category*	*Ref. category*	*Ref. category*	*Ref. category*
Lower secondary degree	0.115	0.111	0.815***	0.180	0.165	0.015
	(0.144)	(0.144)	(0.138)	(0.163)	(0.163)	(0.160)
Higher secondary degree	0.447**	0.444**	1.415***	0.041	0.015	-0.176
	(0.174)	(0.175)	(0.164)	(0.202)	(0.202)	(0.198)
Post-secondary degree	0.704***	0.705***	1.830***	-0.174	-0.219	-0.484**

	(0.211)	(0.211)	(0.199)	(0.275)	(0.275)	(0.270)
Tertiary degree	0.912***	0.912***	2.424***	-0.829***	-0.878***	-1.277***
	(0.235)	(0.235)	(0.217)	(0.295)	(0.294)	(0.285)
Income	-0.049***	-0.048***	-0.048***	-0.045***	-0.047***	-0.047***
	(0.008)	(0.008)	(0.009)	(0.012)	(0.012)	(0.012)
Rural	-0.178***	-0.178***	-0.195***	0.018	0.019	0.035**
	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
Age	-0.021***	-0.021***	-0.020***	-0.009***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Secular	0.212***	0.215***	0.208***	0.142***	0.137***	0.139***
	(0.015)	(0.015)	(0.015)	(0.020)	(0.020)	(0.020)
Country intercept variance	0.446***	0.448***	0.447***	1.737**	1.732**	1.701**
	(0.167)	(0.168)	(0.168)	(0.763)	(0.761)	(0.748)
ISCO intercept variance	0.096***	0.095***	0.178***	0.116***	0.116***	0.146***
	(0.020)	(0.020)	(0.026)	(0.029)	(0.029)	(0.032)
Intercept	-3.113***	-3.161***	-2.313***	2.121***	2.097***	2.980***
	(0.248)	(0.249)	(0.244)	(0.585)	(0.585)	(0.571)
Observations	40,955	40,955	40,955	30,794	30,794	30,794
Groups	15	15	15	11	11	11
Log-likelihood	-11680.3	-11672.8	-11799.5	-6525.3	-6515.9	-6546.5
BIC	23530.4	23526.2	23758.2	13215.9	13207.5	13123.0

E. Replication of the CECT schema in a survey in the United States

We replicate the skills schema developed by van de Werfhorst (2001) for 21 educational fields in a new survey fielded in the United States in March 2023.[1] Each respondent who had at least completed high school was asked to select from a drop-down menu their main subject or field for their highest degree. We take the same subjects as used in the European Social Survey, but disaggregate some to provide separate entries for categories that have expanded over the past decades, such as environmental studies, computing and IT, sports & leasure, food and catering, public administration, and planning. Next, a list of sixteen kinds of skills and knowledge was presented to respondents in random order, and they were asked to evaluate to what extent their education paid attention to these.[2] We used the same wording as in the original study barring some minor stylistic changes, and

[1] This convenience sample was collected in March 2023 by TGM for 800 respondents (IRB 22-0061 at University of North Carolina at Chapel Hill). The survey slightly oversampled Democrats (34.7%) and Independents (22.9%) and undersampled Republicans (26.6%) – 15.8% identified as Democrat- or Republican-leaning Independents, with quotas on age, state, and education.
[2] We follow van de Werfhorst (2001) and our own ESS application in allocating a score of 1 (to a very limited degree) for each of the skills to respondents with less than a high school diploma.

the same five-point scale. Please see the Additional Documentation on Dataverse for question wording in the 2023 US survey.

Reliability analyses show that the cultural, economic, communicative, and technical scales are measured reliably: all four scales have a Cronbach's alpha of 0.76 or higher. A principal components analysis using orthogonal varimax rotation produces four factors with an eigenvalue above 2 that correspond with the four types of skills or knowledge, the one exception being that "writing and reading" loads marginally stronger on communicative than on cultural skills.

We use this individual-level information to create aggregate CECT scores for each field, and find that these scores are correlated 0.84 with the CECT scores estimated from the 1998 Dutch survey. This is noteworthy, not only because of the time lag in the survey but also because the generalist educational system of the United States and the early-track system in the Netherlands stand at opposite ends of a spectrum of educational systems (Strello et al. 2021). We next allocate to each respondent the CECT score that corresponds to their declared main subject of their degree.

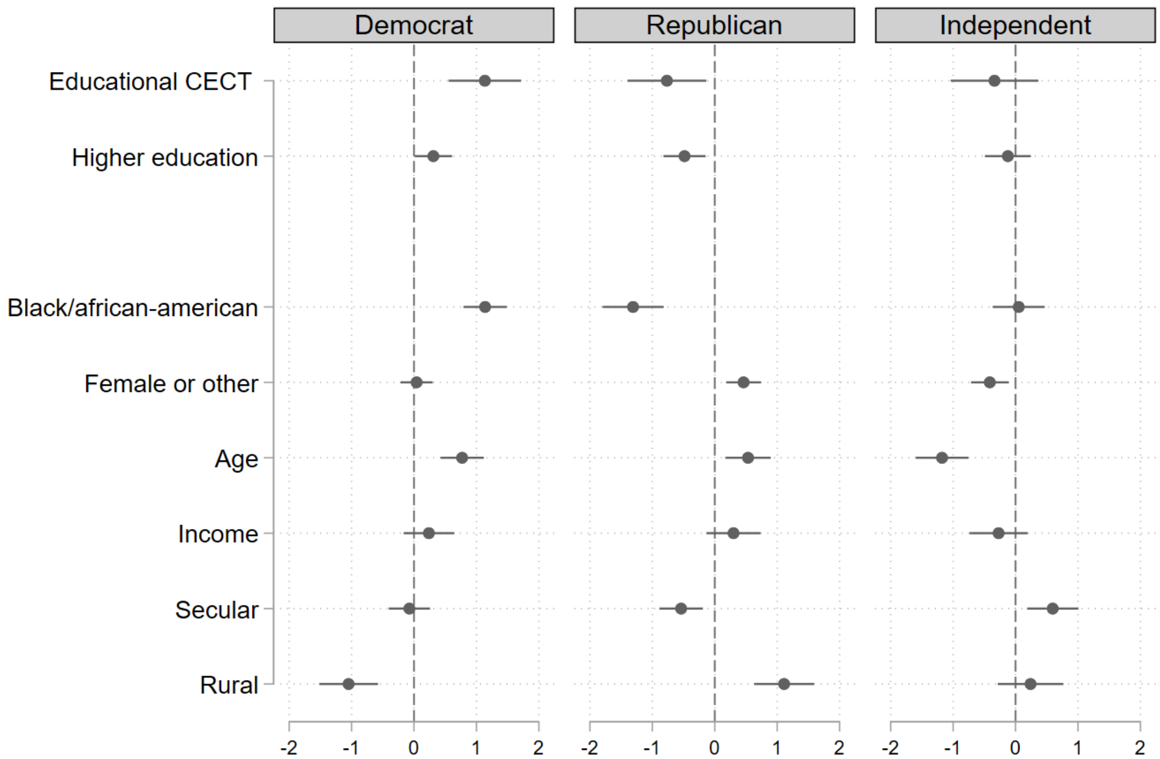
We apply this information to predict strong or weak identification with Democrats, Republicans, or Independents using conventional ANES questions on party identification. Democrat (Republican) takes on a value of 1 if someone identifies strongly or not very strongly as a Democrat (Republican) and zero otherwise; Independent takes on a value of 1 if someone identifies as Independent and closer to neither party, and zero otherwise. Each multivariate logit model contains the same control variables as in the ESS analysis (gender, higher education, rural, age, income, and secular), and we add to this whether a respondent is Black/African-American.

Figure A.1 and Table A.7 shows that field of education is a strong predictor and stronger than level of education. A one-unit increase in educational CECT controlling for all other variables, is associated with an increase in the probability of identifying with the Democratic Party from 25.6% to 49.4% (+/-6.5), and a decrease in the probability of identifying with the Republican Party from 35.2% to 21.3% (+/-6.6). For Democrats, the substantive effect of educational CECT is second only to race, and stronger than rural-urban or religion. For Republicans, it comes third after rural-urban and race.

Level of education matters also, but surprisingly less than field. It is stronger for Republican than Democrat identification: a person with at least a Bachelor's degree has a probability of 22.1% to identify Republican against 30.7% for someone with a lower degree; the corresponding figures are

41% and 34.4% for Democrats. Consistent with our theory that the knowledge or skill content of education constitutes a powerful social marker on the socio-cultural divide, the propensity to identify as an Independent is not significantly shaped by a person's field of education.

Figure A.1. Field of education and party affiliation in the United States in 2023



Note: This figure plots the models in Table A.7, from left to right: explaining whether some identifies Democrat (model 1), Republican (model 2), or Independent (model 3). The coefficients are log odds (with 95% confidence intervals), estimated by a logistic regression.

Table A.7. Field of education and party affiliation using 2023 US CECT data

	Democrat	Republican	Independent
Educational CECT	1.135***	-0.767**	-0.337
	(0.297)	(0.323)	(0.358)
Higher education	0.310**	-0.485***	-0.123
	(0.153)	(0.172)	(0.188)
Female or other	0.042	0.461***	-0.412***
	(0.131)	(0.142)	(0.154)
Black	1.140***	-1.309***	0.052
	(0.177)	(0.250)	(0.212)
Income	0.239	0.300	-0.271
	(0.206)	(0.221)	(0.238)
Secular	-0.074	-0.539***	0.596***
	(0.168)	(0.178)	(0.210)
Age	0.771***	0.533***	-1.177***
	(0.177)	(0.185)	(0.215)
Rural	-1.048***	1.110***	0.240
	(0.240)	(0.245)	(0.269)
Constant	-1.577***	-0.994***	-0.815**
	(0.271)	(0.290)	(0.319)
Pseudo-R^2	0.070	0.083	0.048
Observations	1,123	1,123	1,123

Note: Coefficients are log odds, with standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Educational CECT is continuous; Income, Secular, and Rural are four-category variables; Age is trichotomous; Higher education, Female, and Black are dichotomous. All variables are rescaled 0-1.

F. Panel data from the Dutch LISS and the German SOEP

Table A.8. Key independent variables in LISS

CECT score	A variable ranging from 0 to 1, created using question numbers 11 to 28 from the 'work and schooling' questionnaire. These questions are dummies for what a respondent studied. If someone has studied in a given field, they are coded as having the field-specific score (see van de Werfhorst and Kraaykamp 2001). See the reproduction files for detail.
Later higher educated	A dichotomous variable that takes on a value of 1 if the respondent completed post-secondary non-tertiary education (categories 5 and above of *education*) at any point in their time during the panel, and zero otherwise.
Life phase	Coded based on '*belbezig*', which captures whether someone is in high school, studying, or on the labor market.

Table A.9. Key independent variables in SOEP

CECT score	A variable ranging from 0 to 1, created using several questions (*pgtraina, pgtrainb, pgtrainc, pgtraind, pgisced97*).
Later higher educated	A dichotomous variable that takes on a value of 1 if a respondent at some point in the panel is attending university education or gets their degree.
Attending higher education	Coded based on the life-phase variable, taking on a value of 1 if someone is attending university education and 0 otherwise.
Life phase and in education	Coded based on several variables (*0014, 0267, 0013 (v1 through v3), 0012, 0072*). See the reproduction files for detail.

The main regression model we run for the LISS is as follows:

$$y_{it} = \alpha + laterCECT_i + Postsecondarydegree_i + highschool_{it} + postEdu_{it} + laterCECT_i \times highschool_{it} + laterCECT_i \times postEdu_{it}$$

In the paper, we examine the impact of later CECT over different life phases, and the interaction terms allow us to measure this effect. In the SOEP, we run a similar model, adding a third interaction for the extra life-phase in that model.

Table A.10 shows the full regression model for the forest plot in Figure 6 (right panel), which explains sympathy for GAL (Groenlinks; Party van de Dieren; D66) and TAN (PVV; Forum voor Democratie) parties. Table A.11 shows the full regression model using the SOEP for the forest plot in Figure 6 (left panel), which explains whether respondents lean towards voting for the Green party.

Table A.10. Effect of educational CECT on TAN or GAL vote using the LISS

	DV: TAN thermostat	DV: GAL thermostat
Later educational CECT	-1.236 **	0.707 *
	(0.472)	(0.318)
In Postsecondary or working	-0.561 *	-0.089
	(0.244)	(0.170)
Post-secondary degree completed	-1.039 ***	0.274
	(0.227)	(0.183)
CECT * In Postsecondary or working	-0.001	0.273
	(0.419)	(0.290)
Intercept	5.018 ***	4.838 ***
	(0.283)	(0.183)
R^2	0.078	0.030
Adj. R^2	0.077	0.028
Observations	2801	2748
RMSE	2.507	1.898
N Clusters	444	445

Note: Outcomes are sympathy thermostat variables for TAN parties---PVV and FvD---and GAL parties---GL, D66, and PvdD (0 to 10). We use standard multivariate regression models and coefficients capture changes on the thermostat scale. *** p<.001; ** p<.01; *p<.05.

Table A.11. Effect of educational CECT on party sympathy using the SOEP

	DV: vote intention for the Greens
CECT	0.038 **
	(0.016)
In Post-secondary education	0.000
	(0.007)
Working	-0.017*
	(0.008)
Later higher education	0.065 ***
	(0.006)
Female	0.010
	(0.007)
CECT * In Post-secondary education	0.042 *
	(0.017)
CECT * Working	0.075 ***
	(0.022)
Intercept	-0.006
	(0.007)
R^2	0.032
Adj. R^2	0.032
Observations	46364
RMSE	0.248
N Clusters	4034

Note: The outcome is 'learning Green' (1) or not (0). We use standard multivariate regression models. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

There is one remaining issue with this identification strategy: students in Germany and the Netherlands are tracked in high school. This means that some high school courses prepare students for specific fields in higher education. For example, gymnasium students planning to major in medicine will take biology in high school, while those planning to study economics or engineering usually do not. We believe this is unlikely to affect our results for two reasons. First, even early-track students continue to take a substantial number of generalist courses such as languages or history. Hence early-track students still have significant exposure to different fields. Second, following a track during high school does not make it insurmountable for students to later specialize in a different track. If early tracks are influential, we would expect to find the effect of completed field of study to be smaller.

To test to what extent our results are affected by tracking, Table A.12 repeats the analysis, but now only for respondents for whom we have data in their first or second year in their track in high school, that is, when they are 16 or 17 (as opposed to 18 or older). Hence we focus on people who could only have been exposed to specialized tracking for at most a year or so. We use the larger SOEP sample. If tracking explained our results, we would expect a smaller coefficient for those students compared to those who spent more years in tracking. However, we find the opposite: when restricting the sample to high-school students that have been tracked for fewer years, the effect of later CECT is at least as large as it is for the full sample. This increases our confidence that tracking does not explain the results. Note that due to the interaction terms in the model, the effect of CECT among high-school students is the 'CECT' coefficient in Table A.12.

Table A.12. Effect of educational CECT on vote intention (respondents entering SOEP at 16 or 17)

	DV: vote intention for the Greens
CECT	0.050 **
	(0.018)
In Post-secondary education	-0.006
	(0.008)
Working	-0.010
	(0.009)
Later higher education	0.059 ***
	(0.007)
Female	0.005
	(0.008)
CECT * In Post-secondary education	0.057 **
	(0.020)
CECT * Working	0.062*
	(0.025)
Intercept	-0.002
	(0.008)
R^2	0.036
Adj. R^2	0.036
Observations	33648
RMSE	0.238
N Clusters	2594

Note: The outcome is 'leaning Green' (1) or not (0). We use standard multivariate regression models. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

Table A.13 and A.14 show the full regression models for the results in Figure 8 in the paper. These models aim to capture the effect of occupational CECT while someone is still in education, showing that self-selection into a given occupation explains (part of) the effect of occupational CECT.

Table A.13. The effect of occupational CECT - LISS

	TAN thermostat		GAL thermostat	
	without control for educational CECT	with control for educational CECT	without control for educational CECT	with control for educational CECT
Later occupational CECT	-2.182 ***	-1.346 *	1.135 *	0.334
	(0.538)	(0.610)	(0.486)	(0.521)
In education	0.516	0.571	0.013	-0.027
	(0.302)	(0.300)	(0.235)	(0.233)
Later occupational CECT * In education	0.069	-0.019	0.052	0.118
	(0.597)	(0.595)	(0.484)	(0.482)
Later educational CECT		-0.688 *		0.660 **
		(0.271)		(0.204)
Intercept	3.838 ***	3.777 ***	4.767 ***	4.825 ***
	(0.273)	(0.273)	(0.237)	(0.237)
R^2	0.029	0.034	0.009	0.018
Adj. R^2	0.028	0.033	0.009	0.017
Observations	6631	6625	6497	6491
RMSE	3.569	3.558	2.634	2.620
N Clusters	1084	1081	1081	1078

Note: Outcomes are sympathy thermostat variables for TAN parties---PVV and FvD---and GAL parties---GL, PvdD, and D66 (0 to 10). We use standard multivariate regression models and coefficients capture changes on the thermostat scales. *** p<.001; ** p<.01; *p<.05.

Table A.14. The effect of occupational CECT – SOEP

	Vote intention for the Greens	
	Without control for educational CECT	With control for educational CECT
Later Occupational CECT	0.182 ***	0.025
	(0.054)	(0.061)
In Post-secondary education	-0.059 **	-0.054 **
	(0.020)	(0.020)
Working	-0.079 ***	-0.083 ***
	(0.022)	(0.023)
In Post-secondary education X Occupational CECT	0.212 ***	0.201 ***
	(0.055)	(0.057)
Working X Occupational CECT	0.210 ***	0.250 ***
	(0.061)	(0.064)
Intercept	0.006	0.017
	(0.019)	(0.021)
R^2	0.021	0.030
Adj. R^2	0.021	0.030
Num. obs.	150068	126873
RMSE	0.262	0.257
N Clusters	15483	12163

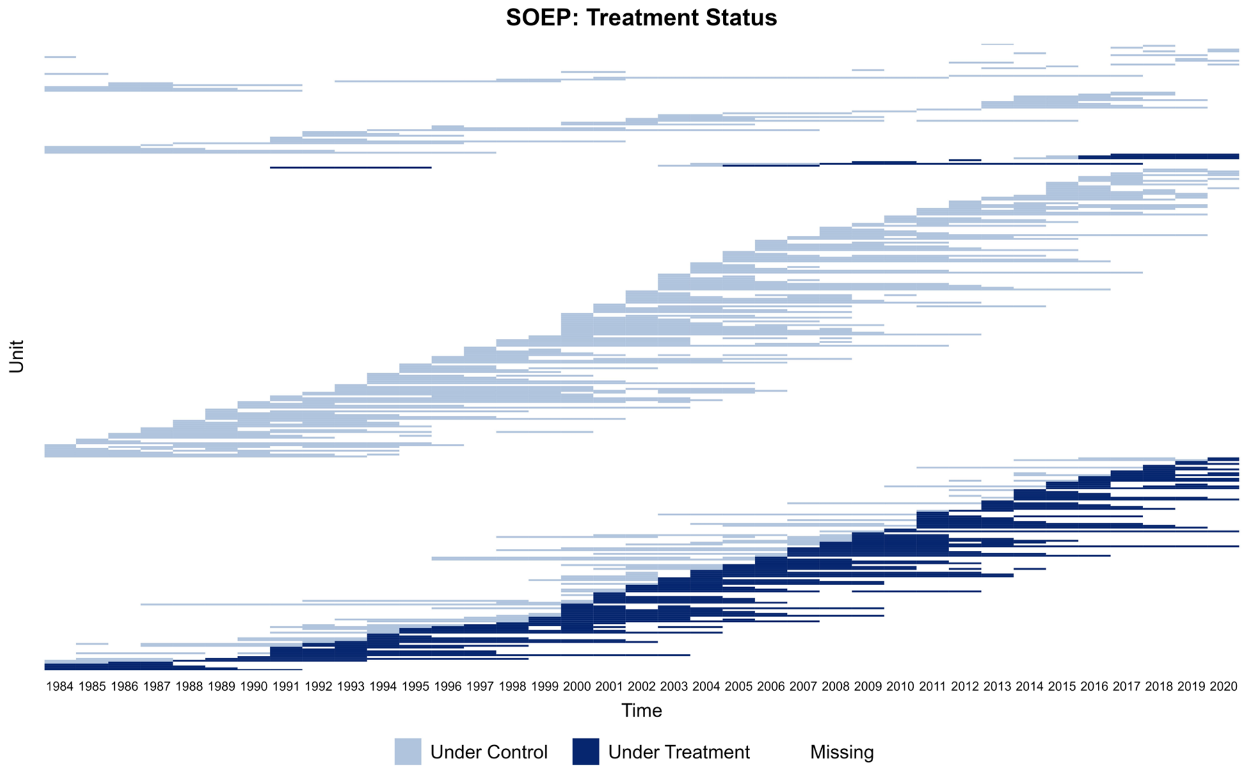
Note: The outcome is 'leaning Green' or not. We use standard multivariate regression models. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

G. Within-individual effects of attending higher education in a CECT field

This section presents modeling results for two alternative approaches for estimating within-individual effects. In the paper (Table 1) we present the results for the 'IFEct' counterfactual estimator developed by Liu, Wang, and Xu (2024). The first model uses a standard Two-Way-Fixed-Effects estimator (Table A.15). The second model uses a Random-Effects-Within-Between Estimator (Table A.16). Figure A.11 visualizes the second and third approach. We use SOEP data for each approach. The results are almost identical using these three different approaches.

Figure A.2 shows the treatment history for a random subset of 500 units (Liu, Wang, and Xu 2024). As expected with individual-level panel data, there is a lot of missingness over the course of the whole panel because few individuals stay in the survey from 1984 up to 2018.

Figure A.2. Treatment history plot



Note: SOEP treatment history for a random subset of 500 units using the '*panelView*' package in *R*.

As we explain in the paper, we prefer the IFEct approach because it does not rely on the assumption of a homogenous treatment effect, which may cause biased estimates, particularly if treatment effects vary depending on when a respondent gets treated (see e.g., de Chaisemartin and D'Haultfoeuille 2020). IFEct does not employ treated observations of early treatment adopters as controls for late treatment adopters, but instead compares each individual to their own counterfactual, and in this way, the estimator accounts for the problems associated with negative weighting in TWFE regressions (Liu, Wang, and Xu 2022). Furthermore, this approach produces more formal plots that allow researchers to assess the parallel trends assumptions (see main paper).

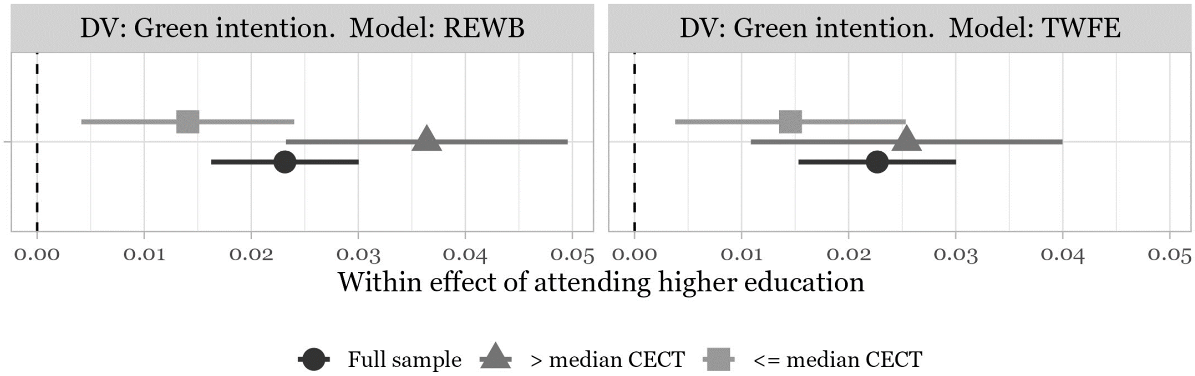
We now show two commonly used alternative approaches to the IFEct approach, one using a Two-Way-Fixed-Effects estimator (Table A.15) and one using a Random-Effects-Within-Between estimator (Table A.16). For each estimator, we apply a generalized Difference in Differences (DiD) approach. A DiD framework estimates the Average Treatment effect on the Treated (ATT).

The TWFE model, until recently the conventional workhorse model for a generalized DiD (also called 'regression DiD'), uses fixed effects for respondents and survey year. Due to the fixed effects, the time-varying treatment variable (attending higher education) is not subject to bias caused by time-constant confounders. However, as several recent studies have shown, the functional form requirements for a TWFE mean that this approach is often biased.

We also present a hybrid approach reliant on a Random Effects Within-Between (REWB) model (Bell et al. 2019), as applied e.g., by Lancee and Sarrasin (2015) and Scott (2022). An REWB model parses out within-individual and between-individual variation, and the latter can be interpreted as the effect of self-selection. Hence by comparing within- and between-effects in the REWB model we obtain an estimate of the relative importance of self-selection versus socialization during or after education. The REWB model also estimates a separate intercept for each individual, which accounts for time-constant confounders, and it estimates a separate beta coefficient on the treatment variable for each respondent, which relaxes the assumption of homogenous treatment effects.

Figure A.3 visualizes REWB (left panel) and TWFE (right panel), and in line with expectations, both approaches find that the within-individual effect of attending higher education is larger for people who graduated in high-CECT fields. Note that the REWB model in Table A.16 estimates the between-individual effect to be about three times as large as the within-individual effect (.073 against .021 in column 1).

Figure A.3. Results from the REWB and TWFE models



Note: SOEP panel using TWFE and REWB models. The effects of higher education are estimated for the full sample, for those with equal to or lower than median CECT, and for those with higher than median CECT. Regression models in Tables A.15 and A.16.

Table A.15. TWFE Within-individual effect of attending higher education in a particular field

	Effect of attending post-secondary education	Attending post sec with > median CECT	Attending post sec with <= median CECT
Attending post-secondary	0.023 *** (0.004)	0.025 *** (0.007)	0.015 ** (0.005)
Reference group used to calculate the common trends	*People in education or working without a post-secondary degree*	*People in education or working without a post-secondary degree and > median CECT*	*People in education or working without a post-secondary degree and <= median CECT*
Individual and time FE	Yes	Yes	Yes
# observations	142558	33120	66866
# respondents	32240	5243	14179

Note: The outcome is 'leaning Green' (1) or not (0). We use a TWFE model to capture within-individual effects. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

Table A.16. REWB Within-individual effect of attending higher education in a particular field

	Effect of attending post-secondary education	Attending post sec with > median CECT	Attending post sec with <= median CECT
Attending post-secondary (within effect)	0.023 ***	0.036 ***	0.014 **
	(0.004)	(0.007)	(0.005)
Attending post-secondary (between effect)	0.094 ***	0.138 ***	0.063 ***
	(0.003)	(0.007)	(0.004)
Intercept	0.054 ***	0.029 *	0.054 ***
	(0.004)	(0.012)	(0.005)
Var respondent intercept	0.025	0.038	0.013
Var within effect	0.03	0.03	0.02
Reference group used to calculate common trends	*People in education or working without a post-secondary degree*	*People in education or working without a post-secondary degree and > median CECT*	*People in education or working without a post-secondary degree and <= median CECT*
# observations	142558	33120	66866
# respondents	32240	5243	14179

Note: The outcome is 'leaning Green' (1) or not (0). We use a REWB model to separately model within-individual and between-individual effects. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

H. The effect of educational field over time

Table A.17 reports a multilevel mixed-effects model with intercepts (random effects) by individual, generation, and survey year. This indicates that the effect of CECT remains important as people age. The data are derived from SOEP (1984-2020). The dependent variable is vote intention for the Greens.

Table A.17. The effect of educational CECT over time

	DV: Vote intention Green
Educational CECT	0.130 ***
	(0.004)
Number of years since 25	-0.000*
	(0.000)
Educational CECT X Number of years since 25	0.000 **
	(0.000)
Intercept	0.007
	(0.009)
# observations	387626
# respondents	49071
BIC	-268358.2
Log Likelihood	134230.6
Num. obs.	387626
Num. groups: syear	37
Num. groups: generation	6
Var: pid (Intercept)	0.027
Var: syear (Intercept)	0.000
Var: generation (Intercept)	0.000
Var: Residual	0.023

Note: The outcome is 'leaning Green' (1) or not (0). We use a multilevel mixed-effects model with random effects by individual, generation, and survey year. The coefficients capture changes in the predicted probability of a respondent leaning Green. *** p<.001; ** p<.01; *p<.05.

I. A robustness test with contemporary educational field data

The paper relies on information from the 2000s, because the European Social Survey uniquely collected information on educational field in 2004, 2006, and 2008. Table A.18 reports multivariate regression results using more recent data from the LISS (collected in 2021, 2022, and 2023) using a model that is nearly identical to our main ESS model. This shows that educational CECT continues to be a highly significant predictor of party sympathy on the socio-cultural divide in the Netherlands. This is the case for each of three plausible operationalizations of the dependent variable: TAN sympathies and Green sympathies (measured in the same way as in the LISS panel analysis).

Table A.18. The effect of field of education in The Netherlands in 2020-22

	DV: TAN Thermostat	DV: GAL Thermostat
Educational CECT	-0.381 ***	0.377 ***
	(0.088)	(0.088)
Higher education	-0.670 ***	0.651 ***
	(0.062)	(0.060)
Female	-0.117 *	0.396 ***
	(0.056)	(0.055)
Income	-0.089 ***	0.000
	(0.024)	(0.027)
Age	-0.024 ***	-0.010 ***
	(0.002)	(0.002)
Migrant background	0.251 ***	-0.000
	(0.076)	(0.067)
Rural	0.056 **	0.092 ***
	(0.018)	(0.018)
Intercept	4.317 ***	3.712 ***
	(0.213)	(0.207)
FE for occupation	Yes	Yes
FE for sector	Yes	Yes
FE for supervising	Yes	Yes
R^2	0.108	0.091
Adj. R^2	0.106	0.090
# observations	6910	6896

Note: OLS model using the latest LISS wave (wave 14, data collection in 2021 and 2022). Rural is a 5-step variable indicating how rural the place is where someone lives. Higher education, female, and migrant background are dummy variables. Age is measured in years and income in 1000 euros monthly net income. As a substitute for someone's ISCO score, we control for occupation, sector, and whether someone is supervising in their job. The outcomes are 0-10 step thermostat scales, and coefficients capture changes on those scales.

References

Bell, Andrew, Malcolm Fairbrother, and Kelvyn Jones. 2019. Fixed and random effects models: making an informed choice. *Quality & Quantity* 53(2): 1051-1074.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110(9): 2964-96.

ESS Round 2: European Social Survey Round 2 Data (2004). Data file edition 3.6. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS2-2004.

ESS Round 3: European Social Survey Round 3 Data (2006). Data file edition 3.7. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS3-2006.

ESS Round 4: European Social Survey Round 4 Data (2008). Data file edition 4.5. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS4-2008.

Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, Jürgen Schupp (2018) The German Socio-Economic Panel Study (SOEP). *Jahrbücher für Nationalökonomie und Statistik/ Journal of Economics and Statistics* (online first), doi: 10.1515/jbnst-2018-0022.

International Labor Organization (ILO). 2012. *International Standard Classification of Occupations: Structure, Group Definitions, and Correspondence Tables*. Geneva: ILO.

Lancee, Bram, and Oriane Sarrasin. 2015. Educated preferences or selection effects? A longitudinal analysis of the impact of educational attainment on attitudes towards immigrants. *European Sociological Review* 31(4): 490-501.

LISS (Longitudinal Internet studies for the Social Sciences) panel administered by Centerdata (Tilburg University, The Netherlands).

Liu, Licheng, Ye Wang, and Yiqing Xu. 2024. "A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data." *American Journal of Political Science* 68(1): 160-76.

Oesch, Daniel. 2006. Coming to grips with a changing class structure. *International Sociology* 21(2): 263-288.

Scott, Ralph. 2022. Does university make you more liberal? Estimating the within-individual effects of higher education on political values. *Electoral Studies* 77: 102471.

SOEP (German Socio-Economic Panel Study (SOEP-Core) Available at DOI (doi:10.5684/soep-core. v37o).

UNESCO. 2014. ISCED *Fields of Education and Training 2013 (ISCED-F 2013). Manual to accompany the International Standard Classification of Education 2011.* Montreal: UNESCO institute for statistics.

Van de Werfhorst, Herman G. and Gerbert Kraaykamp. 2001. Four field-related educational resources and their impact on labor, consumption, and sociopolitical orientation. *Sociology of Education* 74(4): 296-317.

Van de Werfhorst, Herman G. 2001. *Fields of Study and Social Inequality. Four types of Educational Resources in the Process of Stratification in the Netherlands.* PhD Dissertation. Katholieke Universiteit Nijmegen.