Do Parties' Representation Failures Affect Populist Attitudes? Evidence from a Multinational Survey Experiment

–

Online Appendix

Bruno Castanho Silva   Christopher Wratil

Contents

A Issue Questions

Every respondent saw a random draw of two of the following eight issues (Q1–Q8).

Introductory text to all: Thinking about the upcoming European elections in May, we would like to ask your opinion on a few important topics. Please remember there are no right or wrong answers, we are only interested in your opinion. If you feel unsure, please choose the option you feel closer to.

Q1. Some people say that the EU should cut payments to member states accused of violating democratic norms, in order to protect European values. Others say that cutting payments to these countries would be an illegitimate interference with their internal affairs. What do you prefer?

a. Cutting payments.

b. Maintaining payments.

c. Increasing payments.

Q2. Some people say that the EU should have a very strict attitude in the Brexit negotiations, even if it leads to the UK leaving without a deal. Others say that the EU and the UK should achieve a deal even if that means the EU makes large concessions. What do you prefer?

a Making no concessions.

b Making some concessions.

c Making any concessions necessary.

Q3. Some people believe that the EU should take more measures to financially support member states who are in economic trouble. Others think that the support given by the EU today in case of economic crises is already more than enough and should be reduced. What do you prefer?

a Increase financial support.

b Maintain financial support at current levels.

c Reduce financial support.

Q4. Some people say that the EU should keep existing financial support (i.e. subsidies) to farmers in order to protect the existence and quality of European agriculture. Others say that this financial support causes high prices for consumers, and thus subsidies should be cut. What do you prefer?

a Increasing subsidies.

b Maintaining subsidies at current levels.

c Decreasing subsidies.

Q5. Some people think that the EU should impose more restrictions on factory farming to protect animal rights, even if that would increase the price of meat products. Others say that restrictions for factory farming are already too high and should be abolished to lower the price of meat products. What do you prefer?

a Imposing more restrictions.

b Maintaining current restrictions.

c Abolishing some restrictions.

Q6. Some people say the EU should increase military cooperation and build a common European army, which would be stronger and cheaper than national armies. Others say that only the existence of national armies independent from one another can keep countries self-reliant. What do you prefer?

a Increasing military cooperation.

b Keeping current military cooperation.

c Abolishing military cooperation.

Q7. Some people say that the EU should spend more money on social concerns, such as unemployment, even if taxpayers have to pay more into the EU budget. Others say that the EU should reduce spending on social concerns to alleviate the burden on EU taxpayers. What do you prefer?

a Increasing social spending.

b Keeping social spending at current levels.

c Decreasing social spending.

Q8. Some people say that the EU should have no say on how each member state processes asylum-seeking applications. Others say the EU should intervene in asylum-seeking decisions in order to redistribute refugees across Europe. What do you prefer?

a Decreasing EU influence.

b Maintaining current EU influence.

c Increasing EU influence.

B Description of Populism Batteries and Measurement Models

Table A.1: Description and Measurement Model of Pretreatment Populist Attitudes

Statistic	N	Mean	St. Dev.	Populism
Pop1. I would rather be represented by an ordinary citizen than by a specialized politician	22,185	1.783	0.971	1.000
Pop2. The political differences between the elite and the people are greater than the differences among the people	22,081	2.334	1.201	1.480
Pop3. The politicians in the [COUNTRY adj] Parliament need to follow the will of the people	21,181	2.078	1.050	1.274
Pop4. What people call "compromise" in politics is really just selling out on one's principles	21,786	2.511	1.207	1.456
Pop5. The people, and not politicians, should make our most important political decisions	22,206	1.808	0.945	1.254
Pop6. Elected officials talk too much and take too little action	21,357	2.463	1.126	1.318
Pop7. The particular interests of the political class negatively affect the welfare of the people	21,525	1.996	1.019	1.308
Pop8. Politicians always end up agreeing when it comes to protecting their privileges	21,915	1.897	1.038	1.247

Model fit: $chi^2 = 3703.965, df = 305, p < .001$, CFI: .923, TLI: .915, RMSEA: .077 (90% CI: .075–.079), SRMR: .063. N = 22,439.

Notes: **Populism** refers to unstandardized loadings from a one-factor multigroup confirmatory factor analysis, with loadings constrained to be the same across the twelve countries (maximum likelihood robust estimation). Missing data modeled with full information maximum likelihood. Model fit information is respective to that model. All items asked on a 1–5 strongly agree–strongly disagree scale.

Table A.2 contains again the descriptive statistics for the dependent variable of the populism scale, along with a measurement model demonstrating that the items capture the constructs they are supposed to. This is a multigroup Confirmatory Factor Analysis with three factors, one for each theorized dimension of populism (Jöreskog, 1971). Factor loadings are constrained to be the same across the 12 countries – the reason why we report the unstandardized loadings. The model fits the data very well: except for the χ^2 test which is significant, what is expected given the large sample (Kline, 2016), all other model fit indicators are within the recommended ranges for "good" fit by Hu and Bentler (1999). The fact that the model with constrained loadings has good fit to the data, and all loadings are around 1.00, indicates that our item battery is invariant across countries (i.e, differential item functioning is not a significant problem), and can be used in cross-national research (Davidov et al., 2014). Moreover, in a χ^2 test of model difference, which is the most conservative test of measurement invariance, this model is not significantly worse than a *configural* model, in which factor loadings are allowed to vary across groups $(\Delta\chi^2 = 136.89, df = 121, p = 0.15)$.

Table A.2: Descriptive Statistics and Measurement Model for the Populist Attitudes Dependent Variable

	Descriptives			Factor Loadings		
Item	N	Mean	St. Dev.	People	Anti-el.	Manich
Ppl1. Politicians should always listen closely to the problems of the people.	15,599	5.929	1.432	1.000		
Ppl2. Politicians don't have to spend time among ordinary people to do a good job.*	15,480	5.358	1.807	0.720		
Ppl3. The will of the people should be the highest principle in this country's politics.	15,435	5.397	1.523	0.913		
Ant1. The government is pretty much run by a few big interests looking out for themselves.	15,553	5.098	1.709		1.000	
Ant2. Government officials use their power to try to improve people's lives.*	15,425	4.533	1.754		0.731	
Ant3. Quite a few of the people running the government are crooked.	15,536	4.930	1.832		0.958	
Man1. You can tell if a person is good or bad if you know their politics.	15,350	3.132	1.811			1.000
Man2. The people I disagree with politically are not evil.*	15,534	2.854	1.668			0.824
Man3. The people I disagree with politically are just misinformed.	15,630	3.421	1.659			0.775

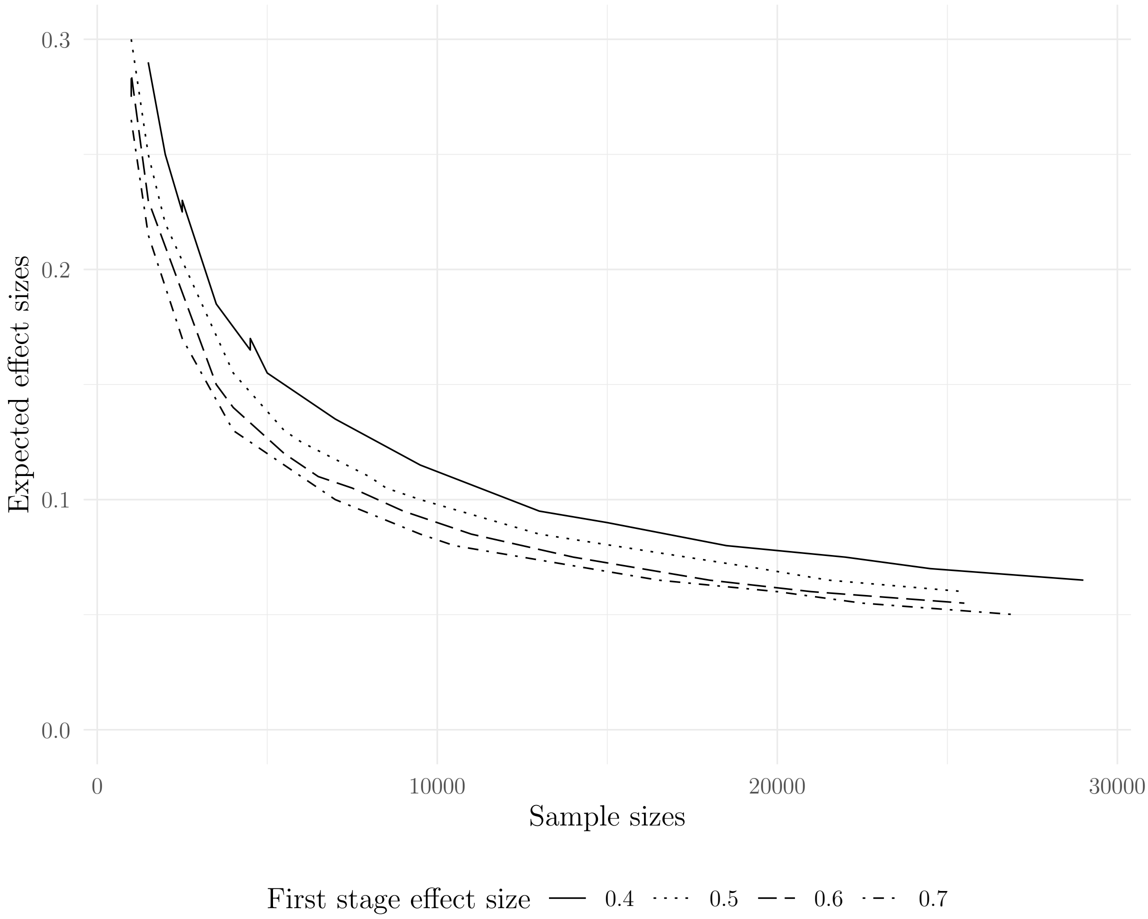
Model fit: $\chi^2 = 1955.148, df = 341, p < .001$, CFI: 0.901, TLI: 0.874, RMSEA: 0.043 (90% CI: 0.041–0.045), SRMR: 0.059. N = 23,257.

Note: All questions asked on a 1–7 strongly disagree–strongly agree scale. Items marked with a * have been reverse-coded to calculate the mean and estimate the factor loading (i.e., for them 1 = strongly agree and 7 = strongly disagree). Factor loadings are unstandardized from a Multigroup Confirmatory Factor Analysis, with loadings constrained to be the same across countries. The model also includes a method factor for the six positive-worded items, with loadings constrained to be the same for all indicators. The method factor is orthogonal to the three substantive factors. Full Information Maximum Likelihood Robust estimation.

C Power Analysis

To test the sample size requirements for our models, we have conducted a power analysis through simulations with the following elements: an instrument Z, which is Bernouilli process with probability of success = 0.42; a confounder W, which is a random normal variable with $\mu = 0$ and $\sigma^2 = 1$; an independent variable X, defined as $X = \beta_1 Z + x*$, in which $x*$ is a normal random variable also with $\sigma^2 = 1, \mu = 0$, and correlated with confounder W at $r = .50$. β_1 is the strength of the relationship between the instrument Z and the independent variable X, to which we attribute four values: 0.4, 0.5, 0.6, and 0.7 (the linetypes in Figure A.1). Y is the dependent variable, predicted by the function $Y = \beta_1 W + \beta_2 X + \epsilon$. When estimating the models to calculate power, we do not include W, so that it is serves as the confounder of the relationship between X and Y which is controlled for, by the IV estimation. We vary β_1 so that it always has the same strength as the first stage (therefore, both are β_1), and β_2 is the effect we are interested in estimating. It varies between 0.01 and 0.3 at 0.005 intervals – these are the y-axis in Figure A.1. For each combination of β_1 and β_2, we use the `ivmodel` package in R (Kang et al., 2020) to estimate the minimal number of responses n required to identify, with a power of 0.9, a β_2 effect significant at $p < .05$ with an instrumental variables model.

Figure A.1: Sample Size Necessary to Identify Expected Effect Sizes of Representation on Populism.



Note: Power analysis conducted with simulated data for an instrumental variable model, varying the effect size of the instrument on the independent variable (line types), the effect size of the independent variable on the dependent variable (y-axis), and the minimal sample size required (x-axis), for those parameters, to have a power of 0.9 of capturing an effect with a significance level of 0.05.

D Descriptive Statistics

Figure A.2: Average Levels of Pretreatment Populism Across Countries

Note: Means and 95% confidence intervals of pretreatment populism, measured as the average response to the Akkerman, Mudde and Zaslove (2014) and Van Hauwaert and Van Kessel (2018) items. Responses range from 1 (least populist) to 5 (most populist).

Table A.3: Sample Descriptive Statistics

	Age	Low Inc.	Med. Inc.	High Inc.	Low Educ.	Med Educ.	High Educ.	Fem.	Left-Right	N
Austria	46	0.17	0.38	0.29	0.27	0.55	0.18	0.47	4.97	1,984
Denmark	52	0.06	0.22	0.51	0.17	0.51	0.32	0.56	5.20	1,973
France	52	0.27	0.36	0.26	0.17	0.50	0.33	0.57	5.23	1,949
Germany	48	0.22	0.36	0.26	0.25	0.49	0.26	0.53	4.74	1,995
Greece	40	0.54	0.25	0.11	0.27	0.22	0.51	0.50	5.04	2,027
Hungary	45	0.72	0.15	0.05	0.14	0.30	0.57	0.52	5.28	1,952
Italy	49	0.47	0.31	0.08	0.45	0.42	0.14	0.56	5.42	1,399
Netherlands	49	0.16	0.40	0.27	0.31	0.37	0.32	0.52	5.41	1,924
Poland	43	0.67	0.20	0.04	0.15	0.62	0.23	0.50	5.31	1,911
Spain	48	0.34	0.36	0.13	0.48	0.25	0.27	0.58	4.46	1,949
Sweden	53	0.12	0.27	0.44	0.11	0.59	0.30	0.56	5.57	1,976
UK	47	0.22	0.29	0.28	0.23	0.38	0.39	0.52	4.87	2,133

Notes: **Age** is the median age; **Low Income** is the share of respondents earning $< 1,500$ EUR/month; **Medium Income** is the share of respondents earning 1,500 EUR to 3,000 EUR/month; **High Income** is the share of respondents earning $> 3,000$ EUR/month; **Left-Right:** mean self-placement on 0–10 Left-right scale.

E Randomization Checks

It is essential to check whether the randomization worked in producing balanced treatment and control groups, so that we can be certain that the effect identified from the experiment is due to the stimulus and not pre-existing differences between the groups. Table A.4 shows the proportion of respondents in the treatment or control groups in each category of education (low, medium, and high) and income group (lower than 1,500 Euros/month, medium, and higher than 3,000 Euros/month), as well as the proportion of female respondents in each treatment condition, and the average age in each one, along with p-values for the mean comparison in each case. Furthermore, since each respondent had two rounds of the experiment where they could be represented or not by a party, we look at how the randomization worked in each of the two issues for each respondent. As the p-values in Table A.4 indicate, the randomzation worked for the entire sample, as there are no significant differences between treated and control (i.e., represented and not-represented) respondents in any of the demographics, across both first and second rounds. This is remarkable particularly given the very large sample size which makes it more likely to identify even small differences as statistically significant. Therefore, we are confident that the results identified are not due to pre-existing differences between the groups.

Table A.4: Randomization Check – Proportion of Respondents in Each Category – Entire Sample

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.241	0.245	0.513	0.247	0.240	0.264
Medium Education	0.435	0.429	0.421	0.432	0.431	0.978
High Education	0.320	0.323	0.717	0.318	0.325	0.260
Low Income	0.320	0.327	0.287	0.328	0.321	0.267
Medium Income	0.287	0.298	0.080	0.292	0.295	0.557
High Income	0.232	0.228	0.482	0.232	0.229	0.587
Female	0.527	0.531	0.568	0.529	0.529	0.943
Age	46.745	47.085	0.108	46.755	47.081	0.123

Note: Values indicate the proportion of the sample belonging to each category for Low/Medium/High Education/Income, proportion of the sample that is Female, and the average age. p denotes a p-value from a two-tailed t-test. *Treated* are respondents whose position was represented by at least one party, while *Control* are respondents whose position was not taken by any party.

In the following tables we report the same tests for each country. While there are eleven significant differences, this is exactly what we would expect by chance given that 192 significance tests are conducted (12 countries, 16 t-tests in each). Moreover, there are no cases where the same variable had an imbalance in a country for both first and second issue.

Table A.5: Randomization Check – Proportion of Respondents in Each Category – Austria

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.270	0.263	0.733	0.274	0.260	0.485
Medium Education	0.548	0.557	0.688	0.545	0.560	0.518
High Education	0.182	0.180	0.899	0.181	0.181	0.974
Low Income	0.165	0.181	0.357	0.172	0.176	0.793
Medium Income	0.353	0.390	0.098	0.380	0.371	0.673
High Income	0.324	0.273	0.015	0.294	0.294	0.999
Female	0.468	0.479	0.639	0.473	0.475	0.906
Age	45.027	45.288	0.730	44.992	45.321	0.659

Table A.6: Randomization Check – Proportion of Respondents in Each Category – Denmark

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.174	0.164	0.558	0.164	0.172	0.643
Medium Education	0.512	0.507	0.830	0.522	0.499	0.301
High Education	0.314	0.329	0.483	0.314	0.329	0.461
Low Income	0.067	0.060	0.516	0.071	0.057	0.222
Medium Income	0.222	0.220	0.917	0.199	0.238	0.040
High Income	0.498	0.521	0.302	0.509	0.512	0.902
Female	0.565	0.563	0.937	0.585	0.548	0.099
Age	50.538	50.303	0.758	49.756	50.905	0.132

Table A.7: Randomization Check – Proportion of Respondents in Each Category – France

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.162	0.178	0.368	0.178	0.167	0.551
Medium Education	0.504	0.499	0.823	0.521	0.487	0.131
High Education	0.334	0.324	0.632	0.301	0.346	0.036
Low Income	0.264	0.271	0.710	0.285	0.257	0.169
Medium Income	0.359	0.353	0.799	0.350	0.359	0.668
High Income	0.252	0.259	0.740	0.256	0.256	0.985
Female	0.563	0.578	0.499	0.575	0.570	0.828
Age	48.993	50.799	0.017	49.594	50.434	0.264

Table A.8: Randomization Check – Proportion of Respondents in Each Category – Germany

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.251	0.246	0.804	0.243	0.252	0.681
Medium Education	0.507	0.475	0.152	0.498	0.483	0.521
High Education	0.242	0.279	0.060	0.259	0.265	0.745
Low Income	0.217	0.213	0.815	0.208	0.220	0.502
Medium Income	0.354	0.366	0.597	0.374	0.351	0.295
High Income	0.251	0.269	0.357	0.257	0.265	0.688
Female	0.534	0.531	0.899	0.537	0.529	0.732
Age	46.112	46.615	0.460	45.986	46.696	0.297

Table A.9: Randomization Check – Proportion of Respondents in Each Category – Greece

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.275	0.269	0.778	0.269	0.273	0.860
Medium Education	0.242	0.210	0.098	0.236	0.214	0.236
High Education	0.484	0.521	0.101	0.495	0.514	0.403
Low Income	0.545	0.529	0.490	0.520	0.546	0.232
Medium Income	0.231	0.267	0.060	0.276	0.236	0.039
High Income	0.127	0.107	0.168	0.114	0.116	0.899
Female	0.507	0.496	0.632	0.472	0.520	0.033
Age	40.271	40.402	0.808	40.859	39.986	0.107

Table A.10: Randomization Check – Proportion of Respondents in Each Category – Hungary

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.122	0.146	0.130	0.138	0.134	0.776
Medium Education	0.299	0.298	0.973	0.298	0.298	0.990
High Education	0.579	0.556	0.315	0.564	0.568	0.853
Low Income	0.711	0.735	0.238	0.708	0.737	0.171
Medium Income	0.169	0.136	0.046	0.155	0.147	0.643
High Income	0.040	0.049	0.304	0.051	0.041	0.292
Female	0.489	0.536	0.042	0.524	0.510	0.564
Age	45.699	46.030	0.643	44.986	46.517	0.032

Table A.11: Randomization Check – Proportion of Respondents in Each Category – Italy

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.412	0.469	0.034	0.461	0.433	0.311
Medium Education	0.436	0.400	0.180	0.415	0.416	0.971
High Education	0.152	0.131	0.267	0.125	0.151	0.156
Low Income	0.457	0.474	0.516	0.493	0.446	0.079
Medium Income	0.308	0.318	0.709	0.289	0.333	0.072
High Income	0.072	0.081	0.497	0.080	0.075	0.701
Female	0.542	0.565	0.383	0.569	0.545	0.373
Age	47.779	49.010	0.125	48.530	48.465	0.936

Table A.12: Randomization Check – Proportion of Respondents in Each Category – Netherlands

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.300	0.319	0.386	0.326	0.300	0.228
Medium Education	0.362	0.373	0.645	0.361	0.373	0.597
High Education	0.338	0.309	0.184	0.312	0.326	0.512
Low Income	0.150	0.170	0.238	0.165	0.160	0.781
Medium Income	0.396	0.396	0.988	0.389	0.401	0.584
High Income	0.280	0.271	0.676	0.284	0.269	0.484
Female	0.511	0.526	0.529	0.534	0.510	0.317
Age	48.819	47.988	0.273	47.771	48.710	0.218

Table A.13: Randomization Check – Proportion of Respondents in Each Category – Poland

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.163	0.138	0.126	0.146	0.150	0.824
Medium Education	0.599	0.637	0.098	0.607	0.632	0.257
High Education	0.237	0.225	0.546	0.247	0.218	0.137
Low Income	0.668	0.675	0.764	0.678	0.668	0.653
Medium Income	0.202	0.200	0.901	0.208	0.196	0.514
High Income	0.051	0.032	0.044	0.037	0.041	0.642
Female	0.519	0.482	0.110	0.478	0.509	0.183
Age	43.843	43.241	0.421	43.265	43.638	0.617

Table A.14: Randomization Check – Proportion of Respondents in Each Category – Spain

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.461	0.494	0.150	0.492	0.472	0.389
Medium Education	0.271	0.236	0.074	0.246	0.254	0.686
High Education	0.268	0.271	0.892	0.263	0.274	0.564
Low Income	0.329	0.344	0.489	0.372	0.312	0.006
Medium Income	0.344	0.366	0.319	0.334	0.374	0.069
High Income	0.125	0.127	0.851	0.130	0.123	0.656
Female	0.569	0.582	0.571	0.584	0.571	0.557
Age	45.564	46.441	0.225	46.229	45.969	0.713

Table A.15: Randomization Check – Proportion of Respondents in Each Category – Sweden

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.106	0.108	0.835	0.108	0.107	0.943
Medium Education	0.601	0.590	0.636	0.581	0.604	0.301
High Education	0.294	0.302	0.713	0.311	0.289	0.290
Low Income	0.121	0.128	0.626	0.124	0.126	0.912
Medium Income	0.254	0.277	0.248	0.277	0.262	0.464
High Income	0.447	0.427	0.377	0.452	0.423	0.215
Female	0.557	0.559	0.903	0.536	0.573	0.103
Age	52.114	52.411	0.700	52.779	51.961	0.286

Table A.16: Randomization Check – Proportion of Respondents in Each Category – UK

	First Issue			Second issue		
	Control	Treated	p	Control	Treated	p
Low Education	0.252	0.221	0.098	0.225	0.240	0.423
Medium Education	0.369	0.385	0.444	0.375	0.381	0.759
High Education	0.379	0.394	0.490	0.400	0.379	0.319
Low Income	0.231	0.212	0.300	0.234	0.208	0.158
Medium Income	0.280	0.299	0.329	0.282	0.298	0.413
High Income	0.272	0.288	0.409	0.271	0.289	0.350
Female	0.531	0.508	0.290	0.522	0.514	0.725
Age	46.491	47.224	0.250	46.910	46.932	0.974

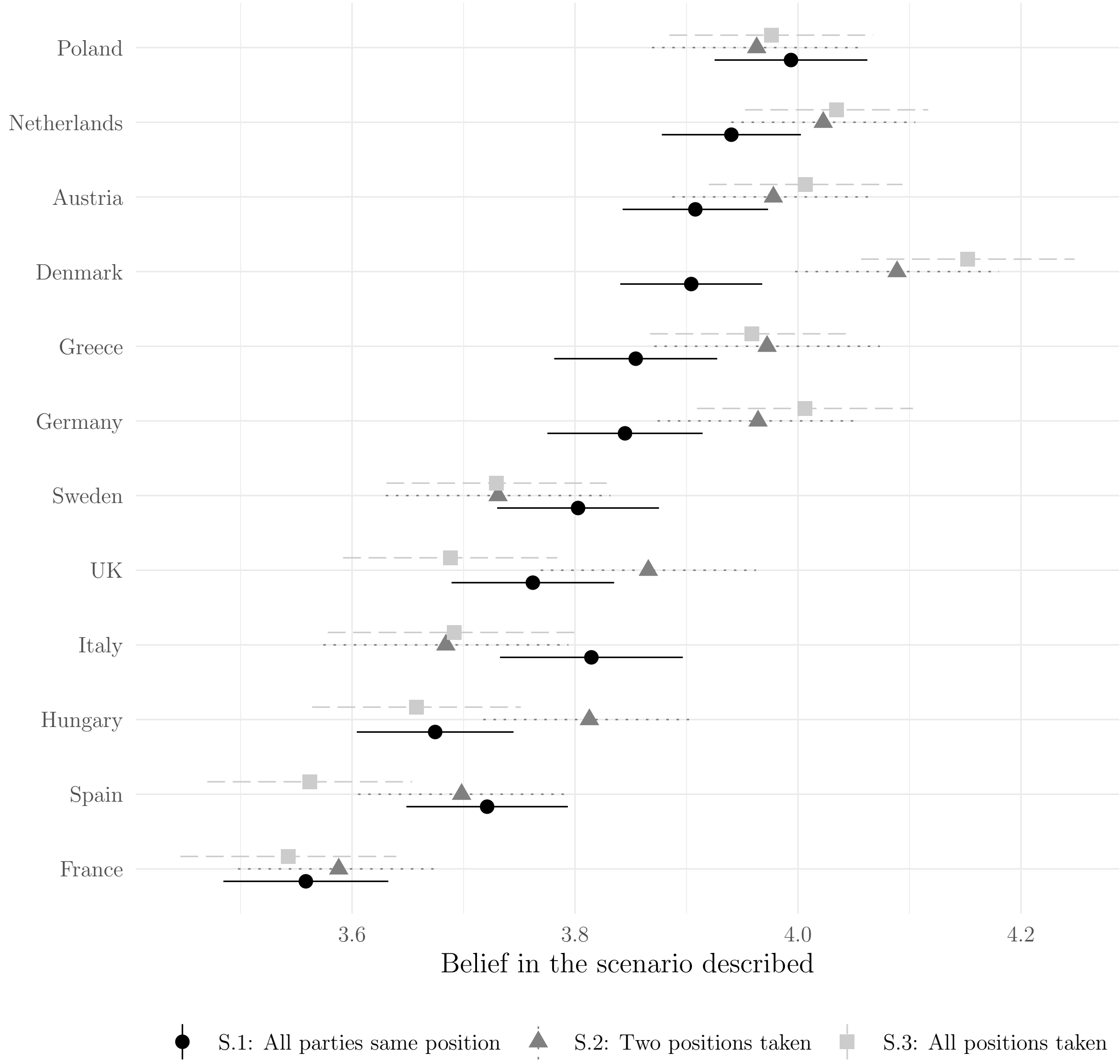
F Manipulation Check

In this section, we analyse our manipulation check for which we asked respondents after each party representation scenario "How likely do you think it is that parties will take the displayed positions?". The results by country in Figure A.3 are described in the paper.

Figure A.4 plots the belief in the described scenarios for the high vs. low populism groups separately. It clearly shows that the low populism group has a higher belief in all scenarios and, in contrast to the high populism group, it identifies the collusion scenario in which all parties take the same position as the least credible one. While the high populism group has a lower belief on average, the absolute value is not so much smaller that we would expect an absence of results purely due to this non-compliance. Moreover, as we see below, the average belief in the treatments by people with low political interest is even lower than the average in the high populism group, while in Figure A.14 below we actually do find significant and comparatively large effects for those with low political interest. Hence, the differences in belief in the scenarios we find are not consistently related to differences in experimental effect sizes. Therefore, while lower belief in the scenarios may contribute to smaller estimates for the high populism group, it is unlikely to be the only or even main cause of the lack of an effect.

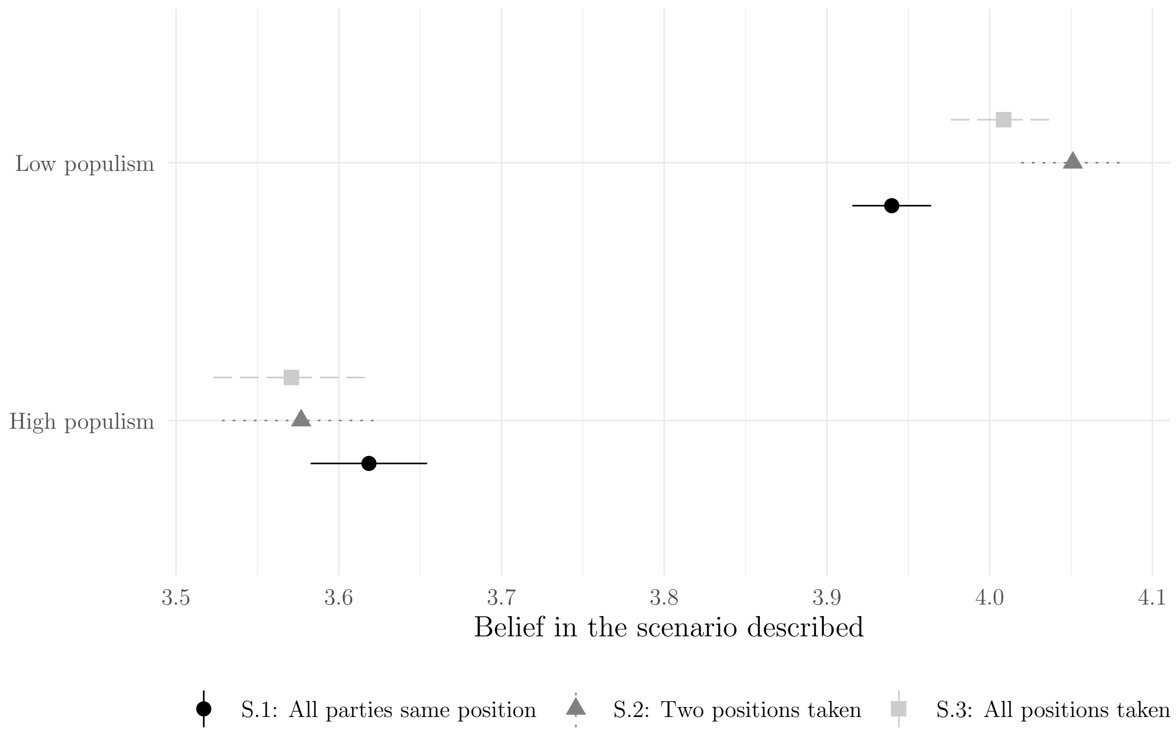
Figure A.5 plots the belief by different levels of political interest. This reveals that the belief in the vignette scenarios increases with political interest, and only individuals with very high political interest are able to identify the collusion scenario as the least likely. The ability to identify the less realistic scenario of party collusion as least likely may contribute to lower experimental effects we observe in the group with high political interest (see Figure A.14 below).

Figure A.3: Belief in the Vignette Scenarios by Country



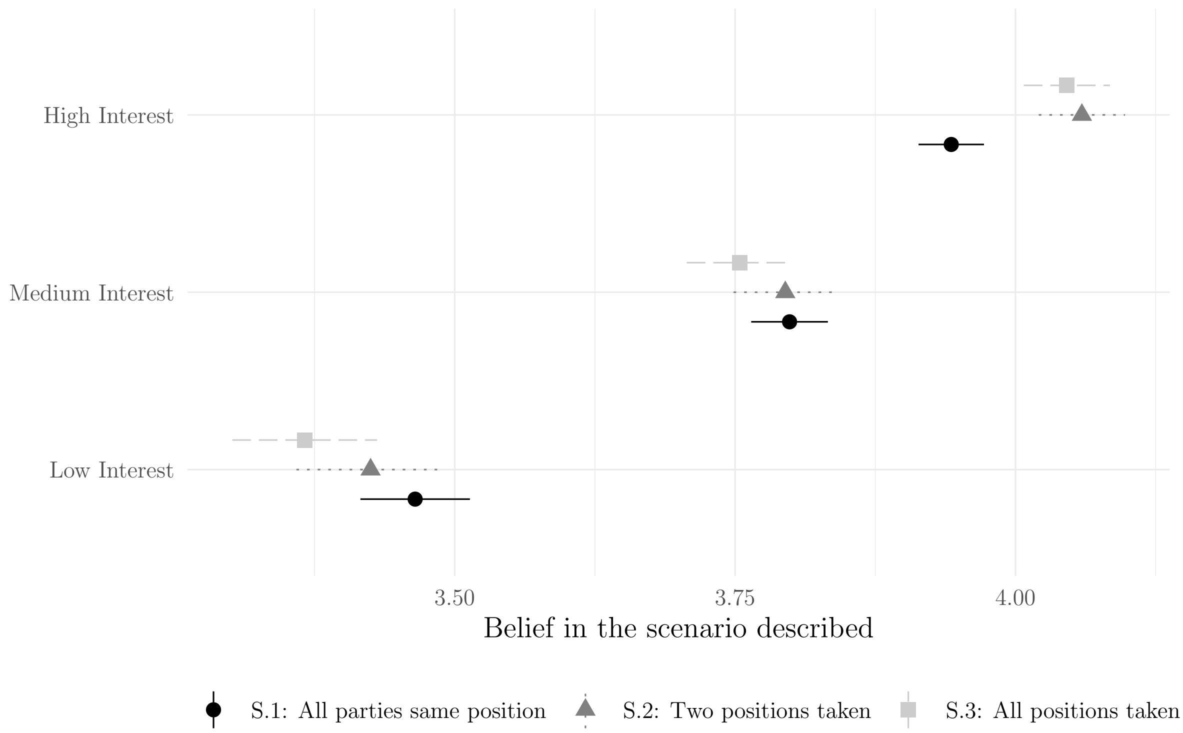
Note: Average answer by country and scenario to the question "How likely do you think it is that parties will take the displayed positions?", with a response scale from 1 (not likely at all) to 7 (very likely), with 95% confidence intervals.

Figure A.4: Belief in the Vignette Scenarios by Pretreatment Populism



Note: Average answer by scenario to the question “How likely do you think it is that parties will take the displayed positions?”, with a response scale from 1 (not likely at all) to 7 (very likely), with 95% confidence intervals; “Low populism” and “High populism” split at the 66^{th} percentile.

Figure A.5: Belief in the Vignette Scenarios by Political Interest



Note: Average answer by scenario to the question "How likely do you think it is that parties will take the displayed positions?", with a response scale from 1 (not likely at all) to 7 (very likely), with 95% confidence intervals; political interest is measured with a 0-10 (not interested at all - very interested) scale, which is grouped into 0-3 as "Low interest", 4-6 as "Medium interest", and 7-10 as "High interest".

G First Stage Estimates

Table A.17: First Stage Estimates for IV Models from Figure 3.

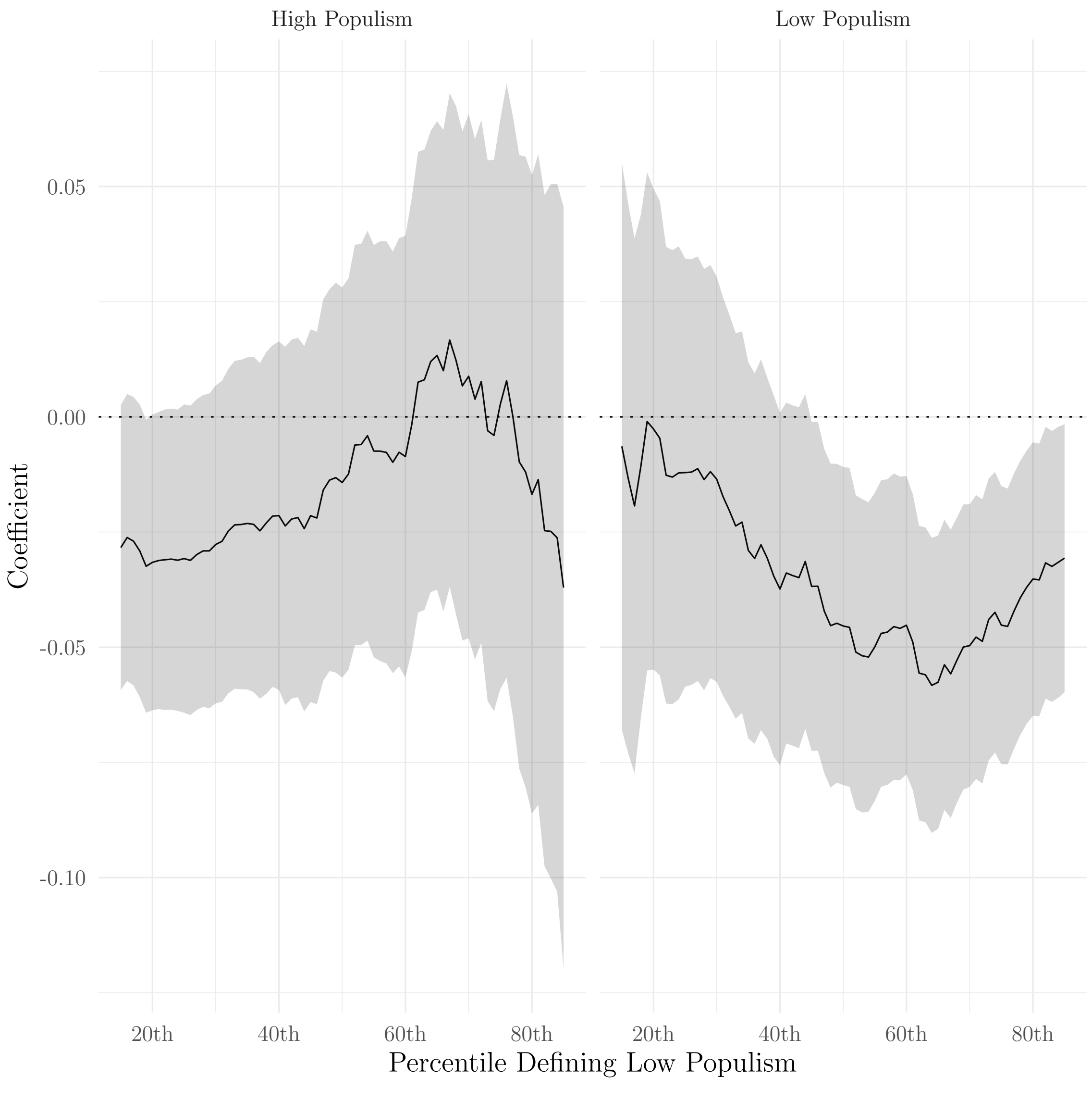
	Full sample	High populism	Low populism
Represented by party	.639 [.610; .667]*	.604 [.553; .654]*	.690 [.655; .724]*
Adj. R^2	.047	.041	.058
Num. obs.	46,513	16,671	28,356

* 0 outside the confidence interval. Dependent variable: feeling well represented by parties.

H Results for Splitting Sample at Different Percentiles

While in the paper we use the 66^{th} percentile for splitting the sample on high vs. low pretreatment populism as a heuristic, we here test the sensitivity of our results by estimating the experimental effects for all splits between the 15^{th} and 85^{th} percentiles. The results are reported in Figure A.6 below. The CACEs for low populism respondents are significant from the 45^{th} percentile onwards, indicating that our results in the paper are no chance finding sensitive to the specific cutoff at which we present the results. There are two potential explanations as to why effects on the low populism sample appear strongest when splitting at the middle of the percentile range, as seen in Figure A.6. The substantive one is that there might be a floor effect, besides the ceiling effect we observe, whereby people very low on populist attitudes are unmoved towards higher populism even if poorly represented. The second option is measurement: studies with information curves from Item Response Theory show that conventionally used populist attitudes scales have less discriminatory power at the lower and upper ends – i.e., for people with very high and very low levels of underlying populist attitudes (Castanho Silva et al., 2019; Van Hauwaert, Schimpf and Azevedo, 2018).

Figure A.6: CACE Estimates Using Different Percentiles to Split the Sample

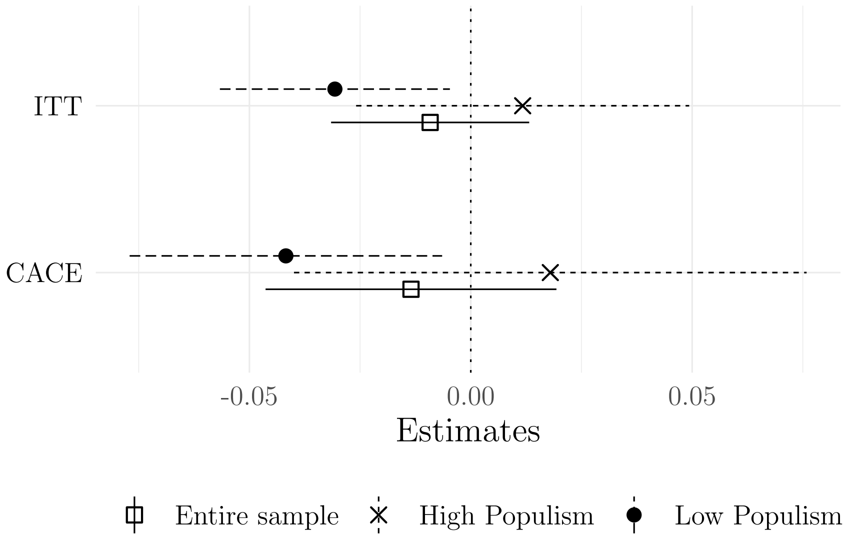


Note: Solid lines denote the regression coefficient of posttreatment populist attitudes on feelings of representation from the split sample instrumental variable models in Figure 3, for different cutoffs to split the sample into high vs. low populism groups. Shaded areas denote the estimated 95% confidence intervals for the models estimated at each percentile cutoff.

I Country-level Results

Recent research has found that populist attitudes scales might not work as expected in countries where populists are in power (Jungkunz, Fahey and Hino, 2021). During the data collection phase, there were four countries in the sample where the largest governing party was considered populist by almost every measure (e.g. the POPPA dataset by Meijers and Zaslove, 2021). These are Italy (Five-Star Movement and Lega coalition), Hungary (Fidesz), Poland (Law and Order), and Greece (SYRIZA). Therefore, we reestimate the models with main effects excluding respondents from those countries. Results in Table A.7 show that effect sizes get a bit smaller but still significant in the same direction for respondents with low pre-treatement populism, and estimates remain substantively the same.

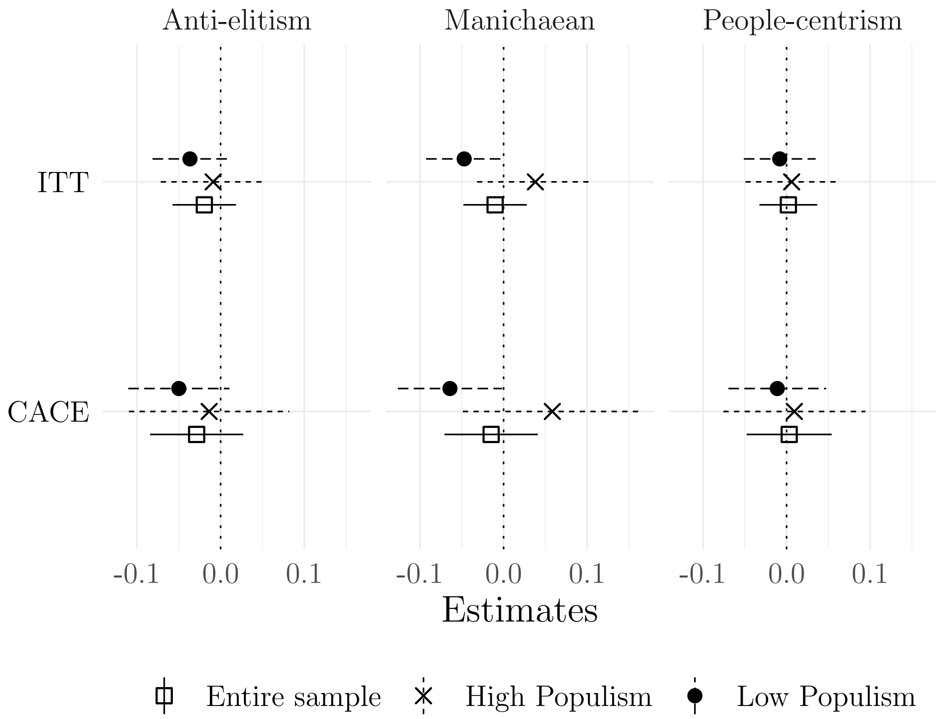
Figure A.7: Effect of Party Representation on Populist Attitudes – Excluding Respondents from Countries with Populist Governments



Note: Dependent variable: posttreatment populist attitudes (mean of three items). **CACE**: *Complier Average Causal Effect* of how well the respondent feels represented by parties on that issue; **ITT**: *Intention-to-Treat Effect* on whether at least one party is shown to hold the respondents' position on an issue. Models include country and round dummy fixed effects. "High vs. Low Populism" defined as above or below the 66^{th} percentile of the pretreatment populist attitudes measure. n (entire sample): 31,765; n (High populism): 11,299; n (Low populism): 19,304. Two rounds per respondents. 95% confidence intervals with robust standard errors.

In Figure A.8 we show these results without countries having populists in power for each subdimension. For the entire sample no estimates are significant, albeit the effect sizes are larger for anti-elitism items, much like for the sample without excluding those countries. For the results by dimension and conditional on pretreatment populism we see that effect sizes remain very similar to those observed in the main analysis, but the effect of anti-elitism among the low populists is not significant anymore due to the reduced sample size, even though the coefficient is almost exactly the same (β = .062 here, and β = .061 with the entire sample including countries with populist governments).

Figure A.8: Effect of Party Representation on Each Dimension of Populist Attitudes – Excluding Respondents from Countries with Populist Governments

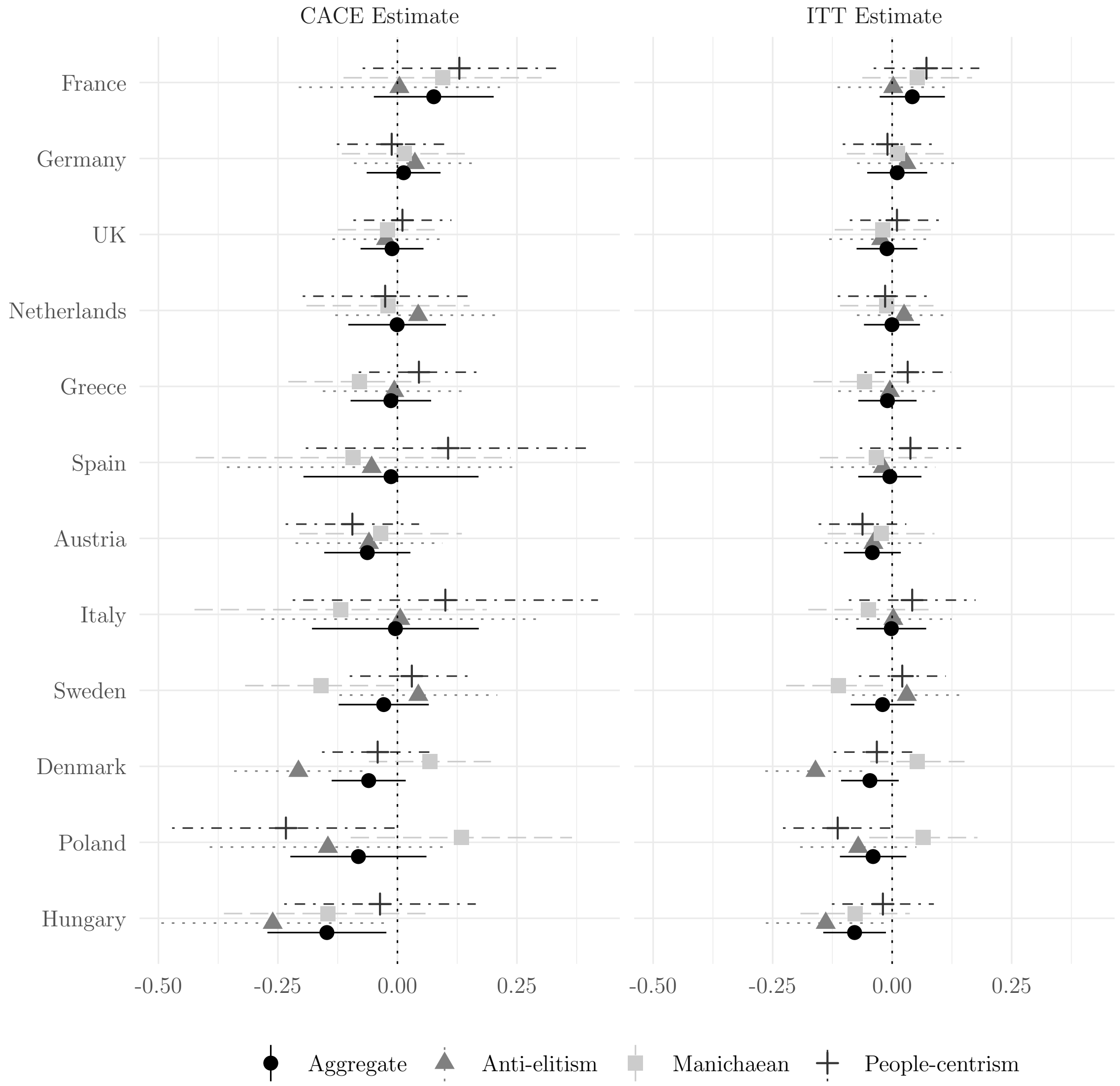


Note: Dependent variables: answer to the relevant item for each dimension of populism. **CACE**: *Complier Average Causal Effect* of how well the respondent feels represented by parties on that issue; **ITT**: *Intention-to-Treat Effect* on whether at least one party is shown to hold the respondents' position on an issue. Models include country and round dummy fixed effects. "High vs. Low Populism" defined as above or below the 66^{th} percentile of the pretreatment populist attitudes measure. n (entire sample): 31,765; n (High populism): 11,299; n (Low populism): 19,304. Two rounds per respondents. 95% confidence intervals with robust standard errors.

In addition, we also report results from running our IV analysis in each country separately. Figure A.9 shows country-specific CACEs and ITTs in the aggregate and for each populism dimension separately. Figure A.10 displays the experimental effects for the low vs. high populism subgroups. Please note that we lack statistical power at the country level, as we only have about 4,000 observations from two experimental rounds for each country. Moreover, for the dimension-specific or subgroup-specific results this sample is further cut in thirds or halves. Even for the aggregate analysis by country, we could only reliably identify the CACE if its size is about 0.1 to 0.15 standard deviations (see section C above) – about two to three times larger than the largest average CACE we find in the low populism group using the entire sample.

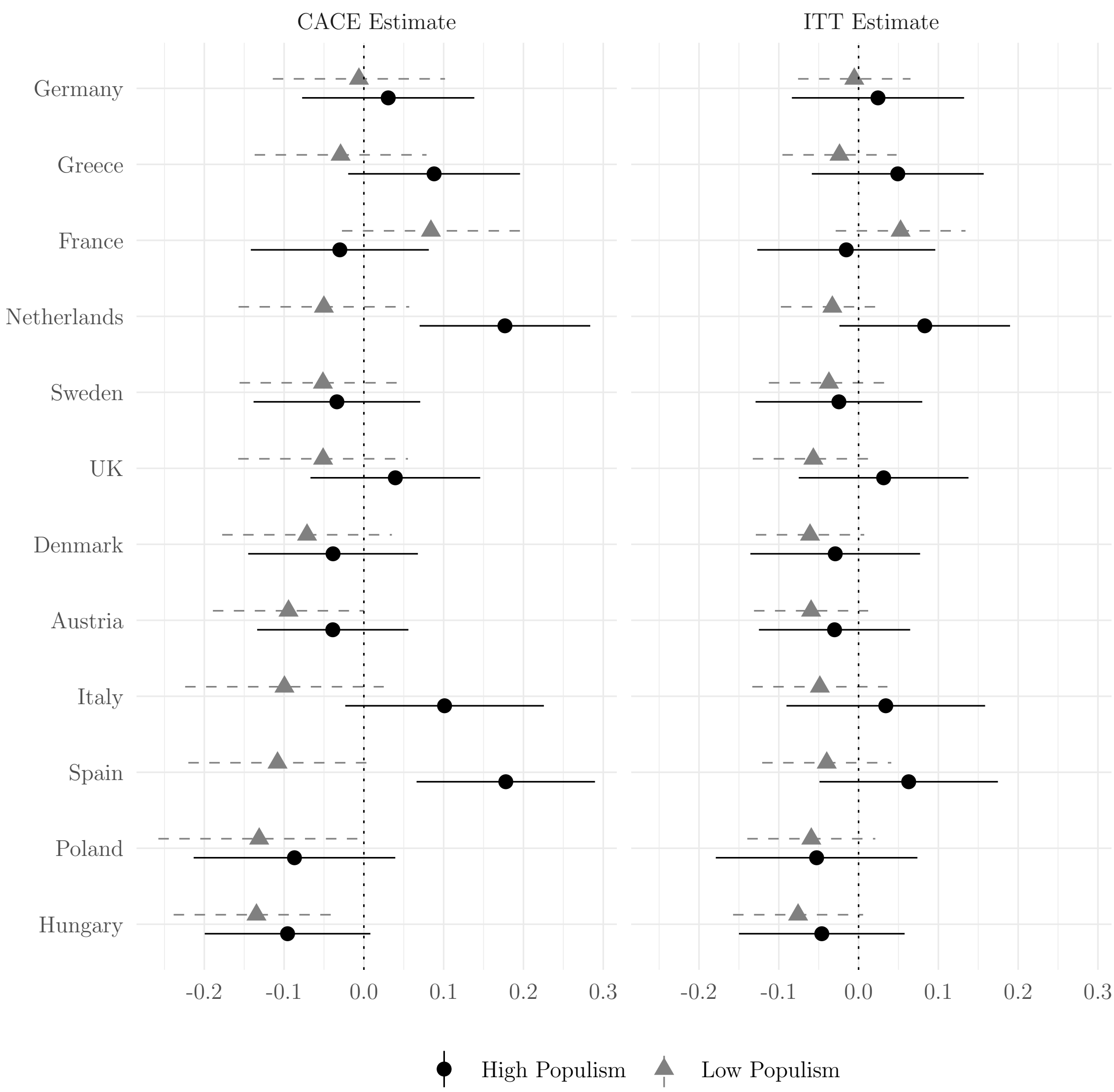
In terms of the robustness of our results, note in particular that 11 of 12 country-level estimates for the low populism group in Figure A.10 are negative, in the expected direction, and several are not far away from statistical significance, albeit the very small sample sizes.

Figure A.9: Experimental Results for the Main Effect and Dimensions by Country



Note: Black dots indicate the results of fitting the models with the full sample in Figure 3 for each individual country separately. The others denote estimates from the models in Figure 4 for each country separately.

Figure A.10: Experimental Results for the Conditional Effect by Country

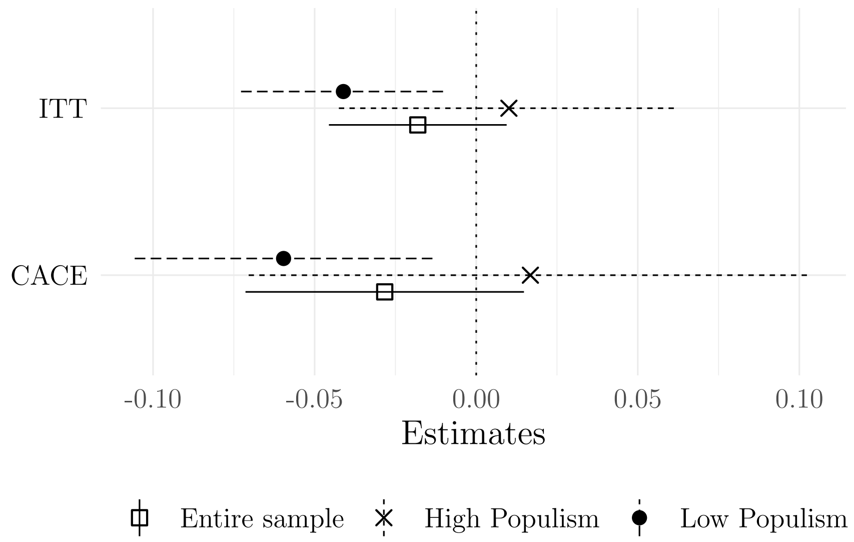


Note: Results from fitting the split sample models in Figure 3 for each country separately, using the same 66^{th} percentile split to define high and low populism groups.

J Alternative Operationalization of Populism

Wuttke, Schimpf and Schoen (2020) propose that, since populist attitudes are multidimensional in a non-compensatory way, a better way to operationalize them is by taking the lowest value a respondent has across each dimension as their level of populism – what they call the "Goertz" operationalization. In this section, we use that operationalization for the posttreatment populist attitudes measure and show that all results remain substantively the same when using it. Note that, on average, effects are marginally weaker with this operationalization, which is expected, since the "Goertz" operationalization only draws on the item with the lowest agreement, thereby essentially estimating effects driven by the Manichaean items (Man1, Man2, and Man3) that have the lowest agreement of all items.

Figure A.11: Effect of Party Representation on Populist Attitudes using Goertz Operationalization



Note: Dependent variable: posttreatment populist attitudes (minimum of three dimensions). **CACE**: *Complier Average Causal Effect* of how well the respondent feels represented by parties on that issue; **ITT**: *Intention-to-Treat Effect* on whether at least one party is shown to hold the respondents' position on an issue. Models include country and round dummy fixed effects. "High vs. Low Populism" defined as above or below the 66^{th} percentile of the pretreatment populist attitudes measure. n (entire sample): 46,513; n (High populism): 16,671; n (Low populism): 28,356. Two rounds per respondents. 95% confidence intervals with robust standard errors.

Figure A.12: Treatment Effects using Different Percentiles to Split the Sample, and Using the Goertz Index of Populism as Dependent Variable.

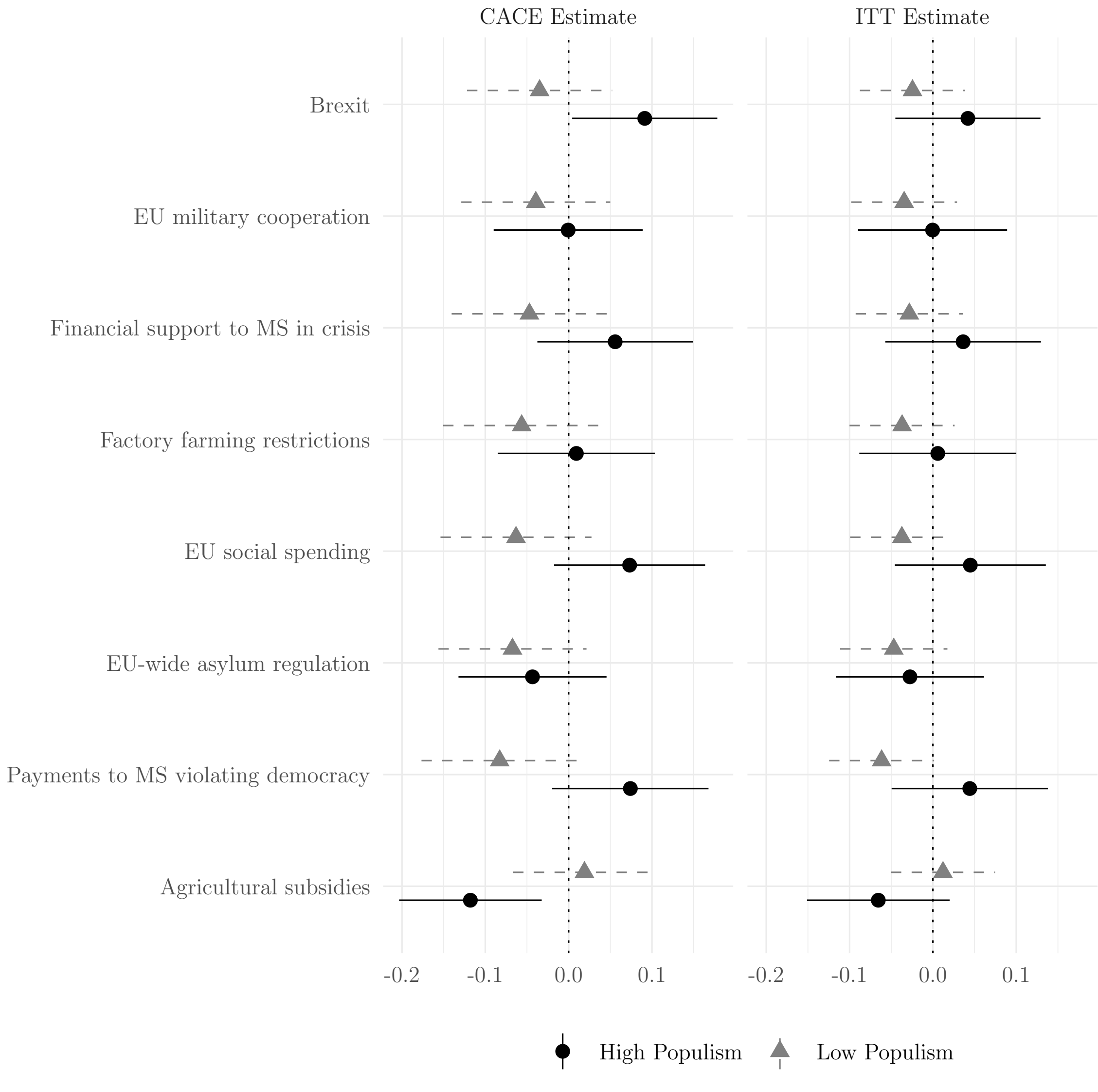
Note: Solid lines denote the regression coefficient of posttreatment populist attitudes on feelings of representation from the split sample instrumental variable models in Figure A.11, for different cutoffs to split the sample into high vs. low populism groups. Shaded areas denote the estimated 95% confidence intervals for the models estimated at each percentile cutoff.

K Results Conditional on Other Factors

In the next pages we investigate the potential effect of other factors on the experimental results. First, we look at whether there are heterogeneous effects by the issue the respondent saw, since some of these topics might elicit different reactions than others. We observe that, for the low populism group, in line with the main findings the effects are negative for all issues except for agricultural subsidies, while for the high populism group effects are a mix of positive and negative. Please note that we have only about 6,000 observations per issue. Hence, we cannot identify our effects for single issues if the expected effect size is about 0.05 standard deviations (the effect we find on aggregate in the low populism group), which would demand a sample closer to 15,000-25,000 observations (see section C above).

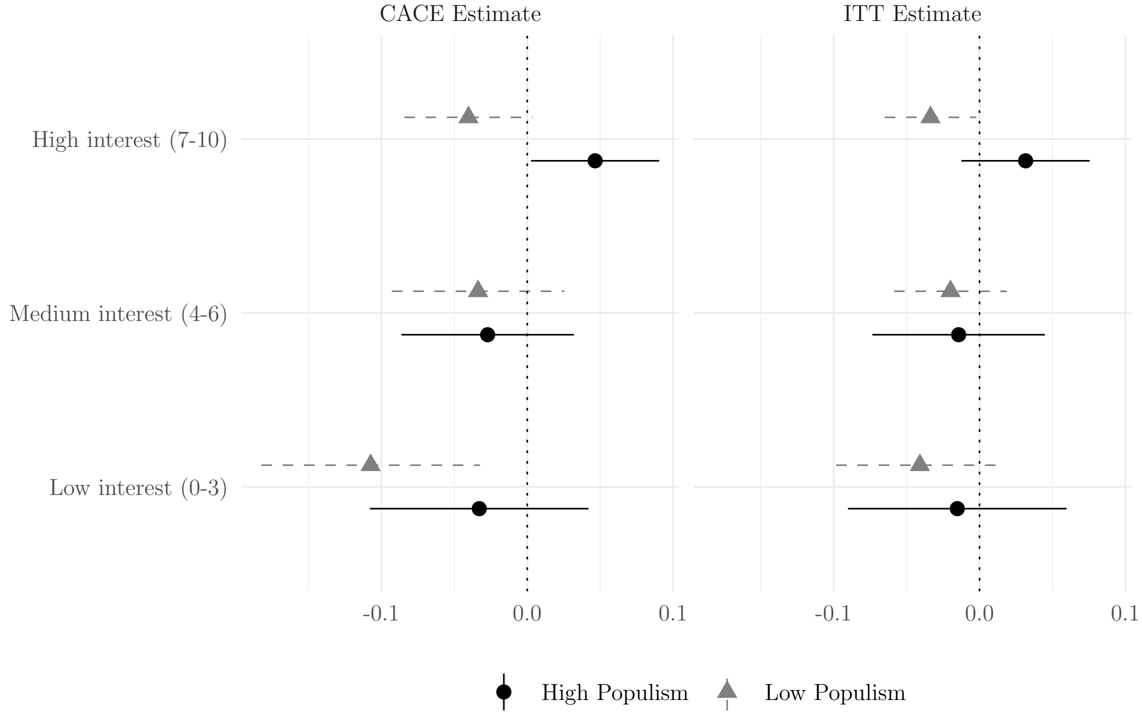
Another possibility is looking into heterogeneous results based on political interest. The design of this experiment presents respondents with hypothetical information about parties which may clash with their perception or knowledge about the political context in the country. It is expected, therefore, that respondents who are more interested in politics may have stronger prior beliefs about how parties would position themselves, and thus would react less to the experimental stimuli. Indeed, this is what we find in Figure A.14: the effect of treatment on individuals low on populist attitudes is stronger and significant for those with low political interest. It is negative but not significant for those with medium and high interest.

Figure A.13: Experimental Results for the Conditional Effect by Issue



Note: Results from fitting the split sample models in Table 3 for each issue separately, using the same 66^{th} percentile split to define high and low populism groups.

Figure A.14: Experimental Results for the Conditional Effect by Pretreatment Political Interest

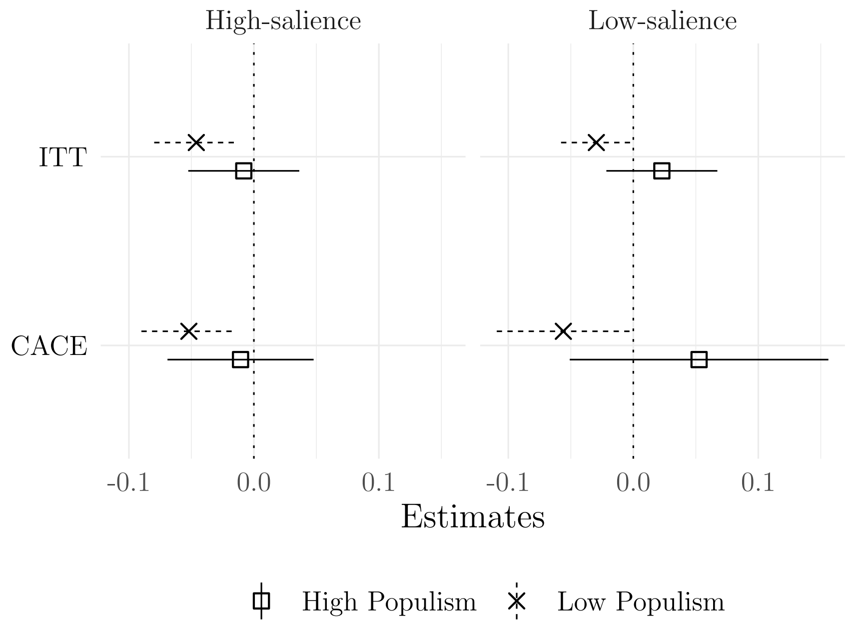


Note: Results from fitting the split sample models in Table 3 for three different levels of political interest separately, using the same 66^{th} percentile split to define high and low populism groups. Political interest is measured with a 0-10 (not interested at all - very interested) scale, which is grouped into 0-3 as "Low interest", 4-6 as "Medium interest", and 7-10 as "High interest".

Finally, we also investigate whether the salience of political issues or dimensions matters for the effects of representation on populist attitudes. If deficiencies in representation occur on policy issues more salient to the individual, such failures may more strongly influence the individual's perceptions and feelings of representation, and such perceptions may also more heavily influence populist attitudes. Directly after asking for their opinion on each issue, we ask respondents how much they care about that issue (on a Likert scale from 1 to 5) in order to tap into issue salience. To split the sample by issue salience, we define as "High Salience" observations in which respondents answered the question on how important the issue was for them with a value above the midpoint of the 5-point scale (i.e., 4 or 5). All other observations are defined as "Low Salience".

Figure A.15 shows the results of the interaction between salience and pretreatment populism levels. This reveals that salience does not appear to matter much once pretreatment populism is factored in: for individuals with low populist attitudes, poor representation leads to higher posttreatment populist attitudes regardless of whether the issue was important to them or not (CACE of −0.052 for the high salience group, and −0.056 for the low salience one). The results suggest that poor representation can trigger populist attitudes for individuals who were not populist beforehand, regardless of the issue in which they are not being represented.

Figure A.15: Effect of Party Representation on Each Dimension of Populist Attitudes by Issue Salience



Note: Dependent variables: posttreatment populist attitudes (mean of three items). **CACE**: *Complier Average Causal Effect* of how well the respondent feels represented by parties on that issue; **ITT**: *Intention-to-Treat Effect* on whether at least one party is shown to hold the respondents' position on an issue. Models include country and round dummy fixed effects. "High vs. Low Populism" defined as above or below the 66^{th} percentile of the pretreatment populist attitudes measure. n (entire sample): 31,765; n (High populism): 11,299; n (Low populism): 19,304. Two rounds per respondents. "High vs Low Salience" defined as an answer above / below (inclusive) the middle category on "how important is this issue for you", ranging from 1 to 5. 95% confidence intervals with robust standard errors.

L Item-level Results

Beyond the possibility of heterogeneous effects across dimensions, which we explore in the paper, it could also be that effects work differently across the different items used to measure populism. While the scales have been subject to extensive development and testing (e.g. Castanho Silva et al., 2018, 2019; Wuttke, Schimpf and Schoen, 2020), there might still be items where the experiment has more impact. Tables A.18 and A.19 show the result of running the models with each item separately as a dependent variable. Note that our statistical power at the level of single items is very limited.

Table A.18: Item-level Experimental Effects for the Entire Sample

Intention-to-Treat Effect Estimates	Ppl1	Ppl2	Ppl3	Ant1	Ant2	Ant3	Man1	Man2	Man3
Being	.03	−.01	−.02	.03	−.07*	−.03	.01	−.04	.01
Represent.	[−.02; .07]	[−.07; .04]	[−.06; .03]	[−.02; .08]	[−.12; −.01]	[−.08; .03]	[−.05; .06]	[−.10; .01]	[−.04; .06]
Adj. R^2	.03	.05	.04	.06	.03	.10	.05	.02	.03
n	15598	15480	15435	15553	15424	15536	15349	15534	15630
Complier Average Causal Effect Estimates									
Feeling	.04	−.02	−.03	.05	−.10*	−.04	.01	−.07	.02
Represent.	[−.03; .11]	[−.10; .06]	[−.11; .06]	[−.04; .14]	[−.19; −.02]	[−.13; .04]	[−.07; .10]	[−.15; .01]	[−.07; .10]
Adj. R^2	.02	.06	.04	.04	.06	.11	.05	.03	.03
n	15598	15480	15435	15553	15424	15536	15349	15534	15630

Notes: * 0 outside the confidence interval. Acronyms for each populism item refer to the statements in Table 2.

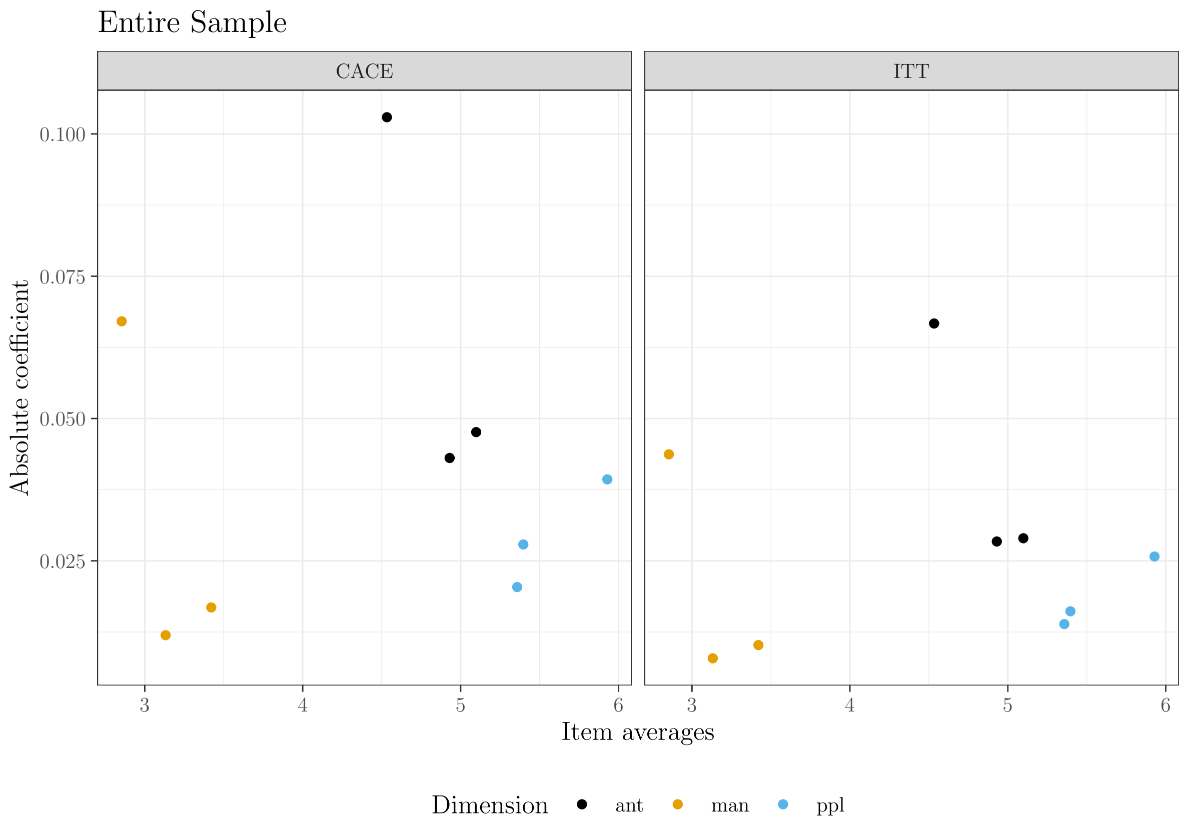
Table A.19: Item-level Experimental Effects by Pretreatment Populism

	Ppl1	Ppl2	Ppl3	Ant1	Ant2	Ant3	Man1	Man2	Man3
Intention-to-Treat Effect Estimates									
					Low pretreatment populism				
Being	.01	−.05	.02	−.02	−.11*	−.04	.00	−.05	−.03
Represent.	[−.05; .07]	[−.11; .02]	[−.04; .08]	[−.08; .05]	[−.17; −.05]	[−.10; .03]	[−.06; .07]	[−.11; .02]	[−.09; .03]
Adj. R^2	.04	.06	.05	.09	.03	.14	.04	.03	.04
n	9500	9435	9421	9501	9441	9414	9447	9420	9489
					High pretreatment populism				
Being	.03	.03	−.04	.07	.00	−.00	−.00	−.05	.07
Represent.	[−.04; .09]	[−.07; .13]	[−.11; .03]	[−.01; .15]	[−.10; .11]	[−.09; .08]	[−.11; .10]	[−.15; .04]	[−.02; .16]
Adj. R^2	.02	.06	.03	.03	.04	.09	.07	.02	.03
n	5600	5546	5525	5563	5484	5624	5413	5621	5637
Complier Average Causal Effect Estimates									
					Low pretreatment populism				
	Ppl1	Ppl2	Ppl3	Ant1	Ant2	Ant3	Man1	Man2	Man3
Feeling.	.02	−.07	.03	−.02	−.16*	−.05	.01	−.07	−.04
Represent.	[−.06; .10]	[−.16; .03]	[−.07; .12]	[−.12; .07]	[−.25; −.07]	[−.15; .04]	[−.09; .10]	[−.17; .02]	[−.14; .05]
Adj. R^2	.04	.06	.05	.09	.07	.15	.04	.03	.03
n	9500	9435	9421	9501	9441	9414	9447	9420	9489
					High pretreatment populism				
Feeling	.05	.04	−.08	.12	.00	−.00	−.00	−.08	.14
Represent.	[−.06; .16]	[−.11; .19]	[−.20; .05]	[−.02; .26]	[−.16; .17]	[−.14; .14]	[−.16; .15]	[−.23; .07]	[−.05; .32]
Adj. R^2	.01	.06	.03	-.01	.04	.09	.07	.02	.03
n	5600	5546	5525	5563	5484	5624	5413	5621	5637

Notes: * 0 outside the confidence interval. Acronyms for each populism item refer to the statements in Table 2.

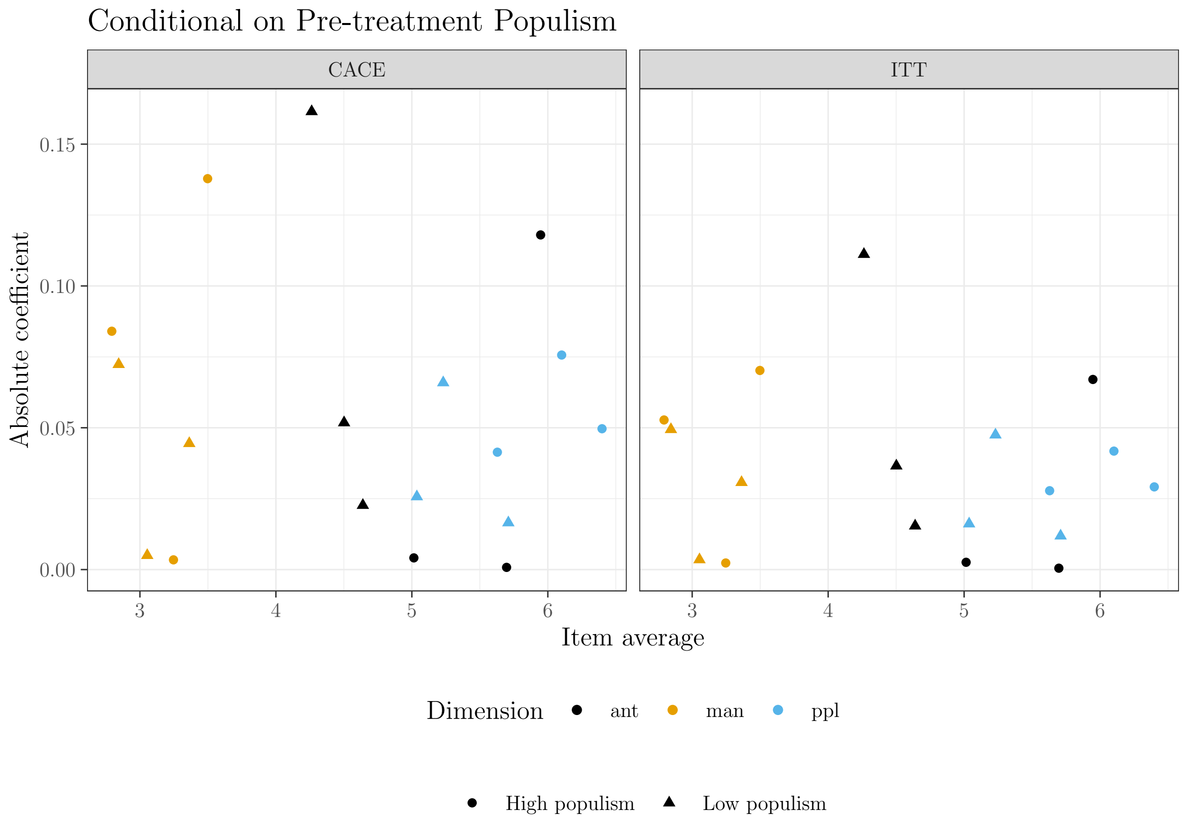
We can also use these estimates to see if there are ceiling effects. The finding that people-centrism is the dimension least affected by the treatment could have a substantive explanation or be a statistical artifact: since it is the dimension with the highest aggregate level of agreement, it could be more difficult to move it. The plots in Figures A.16 and A.17 show the relation between coefficient sizes for each item and the average agreement with the item across the entire sample. If there are ceiling effects, we would expect to see that items with higher averages have lower coefficients. We do not see a clear relationship, as the anti-elitism items with the highest coefficients also have relatively high agreement. Moreover, among the people-centrism items the relationship is even slightly positive, with the most agreeable items yielding the largest coefficients. This suggests that ceiling effects might not play a large role in explaining why people-centrism is not as affected by the treatment as other dimensions.

Figure A.16: Coefficients from Table A.18 and Item Averages - Entire Sample



Note: Absolute size of coefficients from Table A.18 plotted against sample average response to each item, on a 1-7 (strongly disagree – strongly agree) scale.

Figure A.17: Coefficients from Table A.19 and Item Averages - Conditional on Pretreatment Populism

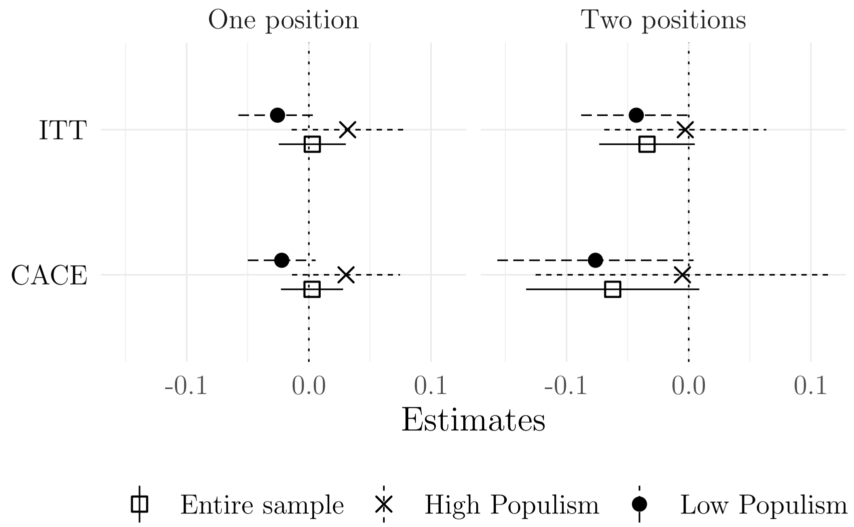


Note: Absolute size of coefficients from Table A.19 plotted against sample average response to each item for the respective group (high or low pretreatment populism), on a 1-7 (strongly disagree – strongly agree) scale.

M Impact of Collective Representation

We check the robustness of our findings on individual representation to different configurations of collective representation. Whereas Scenario 1 in Table 1 represents a situation in which collective representation is poor on average, because only one position is represented by parties, Scenario 3 offers maximal collective representation with every position being represented. Scenario 2 is in-between the two others. To analyze whether our results are dependent on the level of collective representation, we re-estimate the CACE and the ITT effect splitting the sample by vignette scenario (excluding Scenario 3, because all positions are represented in this scenario) as well as by pretreatment populism. The results are in Figure A.18. While none of the effects is statistically significant, given smaller sample sizes, the effect sizes are generally larger for the Scenario 2 sample, in which parties take two of the three positions instead of colluding on a single position. In this scenario, respondents who are not represented see that their position was the *only* one not represented, whereas in Scenario 1 two positions are not represented. Our experimental effects appear largely dependent on rather high levels of collective representation. This speaks against the argument that respondents value collective representation rather than personal representation when considering the scenarios. If collective representation was important we would expect stronger effects if only one position (low collective representation) rather than two positions (medium collective representation) are represented. Our findings point in the exact opposite direction.

Figure A.18: Effect of Representation on Populist Attitudes by Levels of Collective Representation



Note: Dependent variable: posttreatment populist attitudes (minimum of three dimensions). **CACE**: *Complier Average Causal Effect* of how well the respondent feels represented by parties on that issue; **ITT**: *Intention-to-Treat Effect* on whether at least one party is shown to hold the respondents' position on an issue. Models include country and round dummy fixed effects. "High vs. Low Populism" defined as above or below the 66^{th} percentile of the pretreatment populist attitudes measure. n (One position): 23,314; n (Two positions): 11,569. Two rounds per respondents. 95% confidence intervals with robust standard errors.

References

Akkerman, Agnes, Cas Mudde and Andrej Zaslove. 2014. "How Populist Are the People? Measuring Populist Attitudes in Voters." *Comparative Political Studies* 47(9):1324–1353.

Castanho Silva, Bruno, Ioannis Andreadis, Eva Anduiza, Nebojša Blanuša, Yazmin Morlet Corti, Gisela Delfino, Guillem Rico, Saskia Ruth, Bram Spruyt, Marco Steenbergen and Levente Littvay. 2018. Public Opinion Surveys: a New Scale. In *The Ideational Approach to Populism: Theory, Method & Analysis*, ed. Kirk A. Hawkins, Ryan Carlin, Levente Littvay and Cristóbal Rovira Kaltwasser. London: Routledge.

Castanho Silva, Bruno, Sebastian Jungkunz, Marc Helbling and Levente Littvay. 2019. "An Empirical Comparison of Seven Populist Attitudes Scales." *Political Research Quarterly* pp. 1–16.

Davidov, Eldad, Bart Meuleman, Jan Cieciuch, Peter Schmidt and Jaak Billiet. 2014. "Measurement Equivalence in Cross-National Research." *Annual Review of Sociology* 40(1):55–75.

Hu, Li-tze and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6:1–55.

Jöreskog, Karl Gustav. 1971. "Simultaneous Factor Analysis in Several Populations." *Psychometrika* 36:409–26.

Jungkunz, Sebastian, Robert Fahey and Airo Hino. 2021. "Populists Vote for Populists, Right? How Populist Attitude Scales Fail to Capture Support for Populists in Power.".

Kang, Hyunseung, Yang Jiang, Qingyuan Zhao and Dylan Small. 2020. *ivmodel: Statistical Inference and Sensitivity Analysis for Instrumental Variables Model.* R package version 1.8.1.
URL: *https://CRAN.R-project.org/package=ivmodel*

Kline, Rex B. 2016. *Principles and Practice of Structural Equation Modeling.* 4th. ed. New York: The Guilford Press.

Meijers, Maurits J and Andrej Zaslove. 2021. "Measuring populism in political parties: appraisal of a new approach." *Comparative Political Studies* 54(2):372–407.

Van Hauwaert, Steven, Christian Schimpf and Flavio Azevedo. 2018. Public Opinion Surveys. In *The Ideational Approach to Populism: Theory, Method & Analysis*, ed. Kirk A. Hawkins, Ryan Carlin, Levente Littvay and Cristóbal Rovira Kaltwasser. London: Routledge.

Van Hauwaert, Steven M. and Stijn Van Kessel. 2018. "Beyond protest and discontent: A cross-national analysis of the effect of populist attitudes and issue positions on populist party support." *European Journal of Political Research* 57(1):68–92.

Wuttke, Alexander, Christian Schimpf and Harald Schoen. 2020. "When the Whole is Greater Than the Sum of its Parts: On the Conceptualization and Measurement of Populist Attitudes and Other Multi-dimensional Constructs." *American Political Science Review* .