Supplementary Materials

“Fixed Effects and Post-Treatment Bias in Legacy Studies”

Jonathan Homola, Miguel M. Pereira, Margit Tavits

November 2023

These supplementary materials include the following sections:

SM1: Comparison of model specifications

Together, HPT (2020), PGZ, and the current note provide an extensive list of models with different data sources and model specifications. To help the reader follow this collective effort, Table SM1.1 summarizes the different main specifications modeling the effect of camp proximity on contemporary outcomes.

More specifically, we report the results of the main specifications in HPT in Columns 1 (OLS) and 2 (g-estimator). The main results in PGZ are reported in Columns 3 (contemporary state fixed effects) and 4 (Weimar-era states combined with Prussian provinces fixed effects). Finally, the main results of the new model specifications in the current paper are presented across Columns 5 (contemporary state fixed effects only in the first stage of the g-estimator), 6 (Weimar-era state fixed effects in both stages), and 7 (contemporary state fixed effects only in the first stage *and* Weimar-era state fixed effects in both stages). The models with hybrid fixed effects (Weimar states and Prussian provinces) are discussed in SM 5.

Estimates highlighted in green are in line with the theoretical expectations in HPT and reliable at conventional levels. Estimates highlighted in yellow are in line with the theoretical expectations in HPT but not reliable at conventional levels. Across all the different specifications, the only results that go against the original expectations in HPT can be found in Column 3, which present PGZ's models that likely suffer from post-treatment bias as explained in the main text.

Below we describe the estimation equations for the main models reported in the manuscript. The first equation describes a simple OLS model which only contains pre-treatment variables and where i indexes individuals and j indexes states.

$$\text{Exclusionary Attitudes}_{i,j} = \tau_j + \alpha\text{Distance to Camp} + \theta\mathbf{X'}_{i,j} + \epsilon_{i,j}$$

The vector of Weimar-era state fixed effects (τ_j) captures pre-treatment heterogeneity across states at the time of the creation of the camps. The specification also includes a vector of covariates measured pre-treatment that capture pre-existing political attitudes toward out-groups and local economic conditions ($\mathbf{X'}_{i,j}$). The main parameter of interest is α.

The sequential g-estimator, in turn, allows us to consider different contemporary mediators in addition to the pre-treatment variables. The method starts by estimating a model with both pre-treatment and post-treatment covariates in the first stage:

$$\textbf{Stage 1}: \text{Exclusionary Attitudes}_{i,j} = \tau_j + \alpha\text{Distance to Camp} + \theta\mathbf{X'}_{i,j} + \gamma\mathbf{M'}_{i,j} + \epsilon_{i,j}$$

This equation is similar to the OLS specification described above with the addition of the term **M'**, a vector of post-treatment mediators. It is important to note that the purpose of the first stage is to inform the correction that occurs in the second stage. As such, the coefficients for the post-treatment variables should not be interpreted on their own in a substantive fashion.

Next, the estimator recalculates the outcome variable by removing from it the effects of the mediating variables of interest. In the second stage of the g-estimator, we then regress this "demediated" outcome (*Exclusionary Attitudes'*) on the treatment and pre-treatment covariates as follows:

$$\textbf{Stage 2} : \text{Exclusionary Attitudes'}_{i,j} = \tau_j + \alpha \text{Distance to Camp} + \theta \mathbf{X'}_{i,j} + \epsilon_{i,j}$$

Finally, the results from the sequential g-estimator include bootstrapped standard errors to account for the added uncertainty of its two-step nature.

Table SM1.1: Comparing different model specifications for the effect of camp proximity on contemporary outcomes

	HPT 2020		PGZ 2023		HPT 2024		
	OLS	g-estimator	Contemp. FE	Weimar States + Prussian Prov. FE	Contemp. FE	Weimar FE	Contemp. + Weimar FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EVS							
Out-group intolerance	-0.011 (0.003)** [N = 2,075]	-0.017 (0.004)** [N = 1,376]	0.005 (0.005) [N = 1,376]	0.001 (0.005) [N = 1,374]	-0.016 (0.004)** [N = 1,376]	-0.010 (0.004)* [N = 2,055]	-0.018 (0.005)** [N = 1,375]
Immigrant resentment	-0.116 (0.017)** [N = 2,075]	-0.106 (0.020)** [N = 1,376]	-0.018 (0.028) [N = 1,376]	-0.043 (0.030) [N = 1,374]	-0.127 (0.022)** [N = 1,376]	-0.051 (0.023)* [N = 2,055]	-0.097 (0.026)** [N = 1,375]
Far-right support	-0.001 (0.001)* [N = 2,075]	-0.003 (0.001)** [N = 1,376]	-0.001 (0.001) [N = 1,376]	-0.001 (0.001) [N = 1,374]	-0.002 (0.001)* [N = 1,376]	-0.002 (0.001) [N = 2,055]	-0.003 (0.001)* [N = 1,375]
Electoral data							
AfD, full sample		-0.081 (0.016)** [N = 10,755]	0.036 (0.011)** [N = 10,870]	-0.049 (0.010)**† [N = 10,737]	-0.058 (0.015)** [N = 10,755]	-0.087 (0.014)** [N = 10,737]	-0.089 (0.014)** [N = 10,737]
AfD, <70km		-0.159 (0.062)** [N = 3,949]		-0.091 (0.036)*† [N = 3,945]	-0.206 (0.058)** [N = 3,949]	-0.220 (0.045)** [N = 3,945]	-0.038 (0.049) [N = 3,945]
AfD+NPD, full sample		-0.092 (0.017)** [N = 10,755]	0.038 (0.011)** [N = 10,870]	-0.052 (0.010)**† [N = 10,737]	-0.067 (0.016)** [N = 10,755]	-0.094 (0.014)** [N = 10,737]	-0.095 (0.015)** [N = 10,737]
AfD+NPD, <70km		-0.171 (0.066)** [N = 3,949]		-0.096 (0.038)*† [N = 3,945]	-0.220 (0.062)** [N = 3,949]	-0.232 (0.048)** [N = 3,945]	-0.036 (0.052) [N = 3,945]
ALLBUS							
Intolerance (foreigners)	-0.030 (0.013)* [N = 3,081]	-0.047 (0.014)** [N = 2,959]					
Intolerance (Jews)	-0.021 (0.009)* [N = 2,886]	-0.029 (0.010)** [N = 2,787]					
Intolerance (Muslims)	-0.026 (0.012)* [N = 3,233]	-0.041 (0.012)** [N = 3,093]					

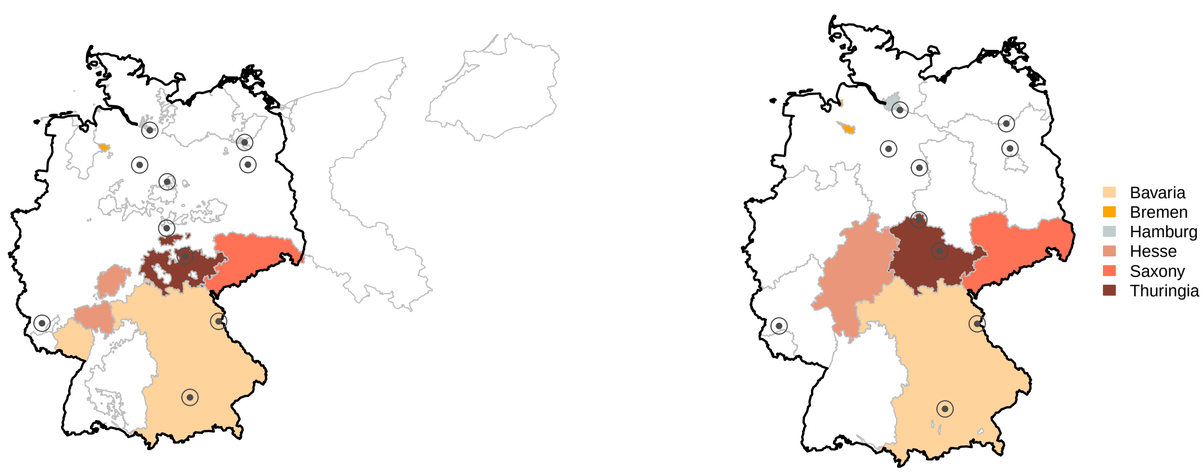
Note: Each row corresponds to an outcome variable, described in the first column. Columns 1 and 2 are based on HPT (2020), Tables 2, 3, and 4. Column 3 refers to PGZ, Table 1, Model 6 for the EVS results and Table 2, Model 2 for the electoral results. Column 4 refers to PGZ, Table A4, Model 6 for the EVS results. Columns 5, 6, and 7 are based on the results presented in the main text and in SM 3 and SM 4. In the electoral data analysis, PGZ do not replicate our results using Weimar FE or subsetting the data to specific radii around the camps. Cells highlighted in green indicate estimates that are in line with our expectations and reliable at conventional levels. Cells highlighted in yellow indicate estimates that are in line with our expectations but not reliable at conventional levels. *p<0.05; **p<0.01; † Results not reported in PGZ.

SM2: Contemporary and Weimar-era states

In Figure SM2.1 we compare the composition of German states in the Weimar era and today. Of the 16 German states that exist today, only six existed under the same name before the first concentration camp was built. And even these six regions changed considerably: for two of them (Hamburg and Hesse), less than 50% of their current territory overlaps with their Weimar-era territory,[1] and only Bavaria remains largely unchanged (although the contemporary state no longer includes the Rhenish Palatinate region).[2] In Western Germany, the collection of states as we know them today has only existed since 1957 when the Saar Protectorate rejoined the Federal Republic as the Saarland. In the GDR, in turn, the states were abolished and only reinstated with new borders in 1990 upon the reunification.

Hence, multiple years – in most cases decades – passed between the construction of the camps and the creation of the German states we know today. Any regional heterogeneity captured by contemporary state-level fixed effects is measured post-treatment and is likely to induce post-treatment bias absent a set of very strong assumptions identified by PGZ.

Figure SM2.1: State borders in the Weimar Republic (1932) and in contemporary Germany



(a) Weimar Republic states **(b)** Contemporary states

Note: Panels (a) and (b) describe the state borders in 1932, the year before the first German concentration camp was created, and in contemporary Germany, respectively. In each panel, gray lines indicate the state borders and the dark line corresponds to the current border of Germany. The shaded states in each panel correspond to the six states from the Weimar period that still exist today under the same name.

[1] In addition, Weimar-era Hesse did not include Frankfurt am Main, which is the state's largest city today.

[2] The proportion of territory in these six states overlapping with their Weimar-era counterparts is: 99.2% (Bavaria), 64.5% (Bremen), 39.0% (Hamburg), 29.7% (Hesse), 79.2% (Saxony), and 70.0% (Thuringia). Additionally, note that once Saxony and Thuringia became part of the GDR, they were abolished and divided up into districts (*Bezirke*) in 1952. This example illustrates the fluidity of regional borders over this period.

SM3: Additional analyses using electoral data

This section presents the full results for the main analysis of the electoral data presented in the manuscript, as well as some additional analyses. More specifically, Tables SM3.1-SM3.3 display the full regression results for the analysis presented in Figure 4.

Matching contemporary geographical units with Weimar-era states is not straightforward. First, contemporary districts are not always contained within a single Weimar-era state. Second, the state of Saarland was not part of Weimar Germany. We therefore use three alternative methods to interpolate Weimar states: (1) matching each contemporary district to the Weimar state that overlaps with the district's geographical center (*centroid interpolation*), (2) using the same centroid interpolation but including Saarland respondents/districts as an additional Weimar state, and (3) matching each contemporary district to the Weimar state that overlaps with the largest share of its area (*area interpolation*). The analyses reported in the main text only included the first interpolation method. However, Figure SM3.1 shows that the findings are also robust to using the other matching approaches described (i.e., centroid interpolation while also including Saarland respondents, and area interpolation).

Table SM3.1: The controlled direct effect of camp proximity on support for radical right parties in 2017, with contemporary state fixed effects in first stage of g-estimator

	AfD Vote Share		AfD + NPD Vote Share	
	Full sample	$< 70km$	Full sample	$< 70km$
	(1)	(2)	(3)	(4)
Distance	-0.058^{**}	-0.206^{**}	-0.067^{**}	-0.220^{**}
(in 10kms)	(0.015)	(0.058)	(0.016)	(0.062)
Nazi party share (1933)	0.008	-0.004	0.012^{*}	0.0004
	(0.005)	(0.010)	(0.006)	(0.010)
% Unemployed (1933)	0.168^{**}	0.120^{**}	0.177^{**}	0.129^{**}
	(0.028)	(0.033)	(0.029)	(0.036)
Population (1925)	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00001)	(0.00000)	(0.00001)
% Jews (1925)	-2.969^{**}	-4.306^{**}	-3.144^{**}	-4.604^{**}
	(0.147)	(0.459)	(0.154)	(0.485)
Current state FEs (N=16)	✓	✓	✓	✓
Contemporary variables	✓	✓	✓	✓
Observations	10,755	3,949	10,755	3,949
Adjusted R^2	0.058	0.084	0.059	0.086

Note: Entries are coefficients of the controlled direct effect of distance to closest camp on support for the AfD (Columns 1-2) and AfD+NPD (Column 3-4) in 2017, corresponding to Table 4 in HPT. All models report the second stage of the sequential g-estimator (bootstrapped standard errors in parentheses). All models include contemporary state fixed effects and contemporary mediators and confounders in the *first* stage regression. *p<0.05; **p<0.01

Table SM3.2: The controlled direct effect of camp proximity on support for radical right parties in 2017, accounting for systematic differences across Weimar states (interpolated from centroids of contemporary Gemeinden)

	AfD Vote Share		AfD + NPD Vote Share	
	Full sample	$< 70km$	Full sample	$< 70km$
	(1)	(2)	(3)	(4)
Distance	-0.087^{**}	-0.220^{**}	-0.094^{**}	-0.232^{**}
(in 10kms)	(0.014)	(0.045)	(0.014)	(0.048)
Nazi party share (1933)	-0.028^{**}	-0.039^{**}	-0.028^{**}	-0.040^{**}
	(0.004)	(0.008)	(0.004)	(0.009)
% Unemployed (1933)	0.055^{**}	0.027	0.054^{**}	0.028
	(0.015)	(0.019)	(0.016)	(0.021)
Population (1925)	−0.00000	0.00001	−0.00000	0.00001
	(0.00000)	(0.00001)	(0.00000)	(0.00001)
% Jews (1925)	-0.505^{**}	-1.001^{**}	-0.519^{**}	-1.105^{**}
	(0.119)	(0.256)	(0.124)	(0.269)
Weimar state FEs (N=17)	✓	✓	✓	✓
Contemporary variables	✓	✓	✓	✓
Observations	10,737	3,945	10,737	3,945
Adjusted R^2	0.390	0.361	0.389	0.364

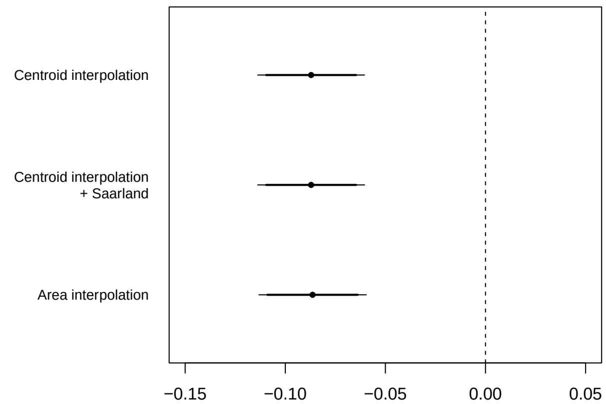
Note: Entries are coefficients of the controlled direct effect of distance to closest camp on support for the AfD (Columns 1-2) and AfD+NPD (Column 3-4) in 2017, corresponding to Table 4 in HPT. All models report the second stage of the sequential g-estimator (bootstrapped standard errors in parentheses). All models include Weimar state fixed effects (interpolated based on the centroids of contemporary Gemeinden) in *both* stages and contemporary mediators and confounders in the *first* stage regression. *p<0.05; **p<0.01

Table SM3.3: The controlled direct effect of camp proximity on support for radical right parties in 2017, with current state fixed effects in first stage of g-estimator and accounting for systematic differences across Weimar states (interpolated from centroids of contemporary Gemeinden)

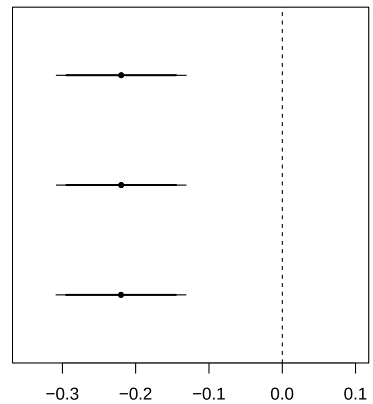
	AfD Vote Share		AfD + NPD Vote Share	
	Full sample	$< 70km$	Full sample	$< 70km$
	(1)	(2)	(3)	(4)
Distance	-0.089^{**}	-0.038	-0.095^{**}	-0.036
(in 10kms)	(0.014)	(0.049)	(0.015)	(0.052)
Nazi party share (1933)	0.009^{*}	0.006	0.012^{**}	0.009
	(0.004)	(0.009)	(0.004)	(0.010)
% Unemployed (1933)	0.080^{**}	0.085^{*}	0.081^{**}	0.091^{*}
	(0.018)	(0.033)	(0.019)	(0.036)
Population (1925)	-0.00000	0.00000	-0.00000^{*}	0.00000
	(0.00000)	(0.00001)	(0.00000)	(0.00001)
% Jews (1925)	-1.088^{**}	-1.914^{**}	-1.150^{**}	-2.089^{**}
	(0.127)	(0.350)	(0.132)	(0.370)
Current state FEs (N=16)	✓	✓	✓	✓
Weimar state FEs (N=17)	✓	✓	✓	✓
Contemporary variables	✓	✓	✓	✓
Observations	10,737	3,945	10,737	3,945
Adjusted R^2	0.415	0.410	0.416	0.417

Note: Entries are coefficients of the controlled direct effect of distance to closest camp on support for the AfD (Columns 1-2) and AfD+NPD (Column 3-4) in 2017, corresponding to Table 4 in HPT. All models report the second stage of the sequential g-estimator (bootstrapped standard errors in parentheses). All models include contemporary state fixed effects and contemporary mediators and confounders in the *first* stage regression. All models also include Weimar state fixed effects (interpolated based on the centroids of contemporary Gemeinden) in *both* stages. *p<0.05; **p<0.01

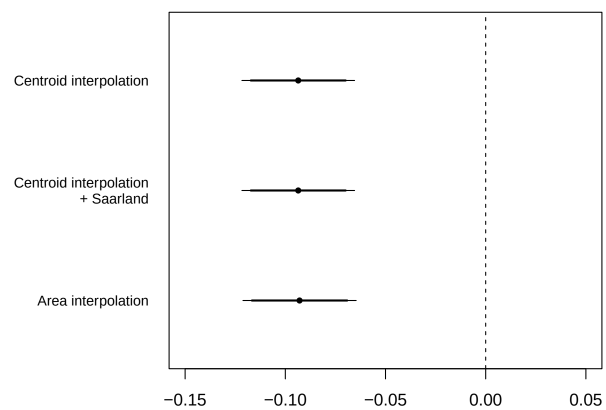
Figure SM3.1: The controlled direct effect of camp proximity on support for radical right parties in 2017, with alternative methods to interpolate Weimar states from contemporary Gemeinden



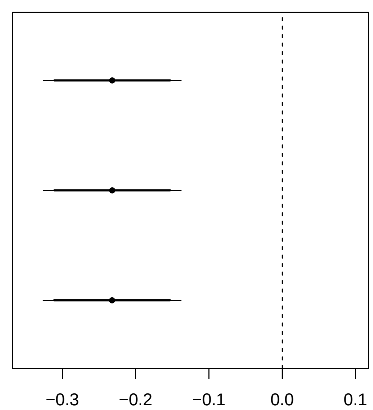
(a) AfD, full sample



(b) AfD, 70km radius



(c) AfD+NPD, full sample



(d) AfD+NPD, 70km radius

Note: Plots depict estimates and 95/90% confidence intervals from the sequential g-estimator for the controlled direct effects of distance to camps on support for radical right parties in 2017 (described in each panel label). Each estimate corresponds to a different model based on alternative methods to match contemporary districts with Weimar-era states. *Centroid interpolation* estimates come from Table SM3.2, *Centroid interpolation + Saarland* estimates from Table DA.1 in the Dataverse Appendix, and *Area interpolation* estimates from Table DA.2.

SM4: Additional analyses using EVS data

This section presents supplementary analyses of the EVS data. More specifically, Tables SM4.1-SM4.3 display the full regression results for the analysis presented in Figure 5. In the main text, we only report results with *centroid interpolation*. However, Figure SM4.1 shows that the findings are also robust to using the other matching approaches described in the previous section (i.e., centroid interpolation while also including Saarland respondents and area interpolation). Finally, while these main models are all replications of our g-estimation approach, the results in Figure SM4.2 show that the "pre-treatment only" OLS models are also robust to the properly specified inclusion of state fixed effects.

Table SM4.1: The controlled direct effect of camp proximity on contemporary attitudes, with current state fixed effects in first stage of g-estimator

	Outgroup Intolerance		Immigrant Resentment		Support Far-Right Parties	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to camp	0.005	−0.016**	−0.023	−0.127**	−0.0003	−0.002*
(in 10kms)	(0.005)	(0.004)	(0.029)	(0.022)	(0.001)	(0.001)
% Jews (1925)	7.561	−0.782	−5.063	0.185	1.891	0.336
	(6.289)	(1.609)	(35.980)	(11.344)	(1.564)	(0.454)
% Unemployed (1933)	1.321	2.177*	1.837	12.369*	0.266	0.299
	(0.909)	(0.934)	(5.202)	(5.694)	(0.226)	(0.195)
Population (1925)	−0.017	−0.013	−0.108	−0.104	−0.006	−0.0004
	(0.016)	(0.012)	(0.093)	(0.077)	(0.004)	(0.003)
Nazi party share (1933)	−0.726**	−0.244	−5.109**	−3.705*	−0.170*	−0.078
	(0.280)	(0.242)	(1.605)	(1.696)	(0.070)	(0.063)
Contemporary covariates						
Conservatism	0.049**		0.292**		0.039**	
	(0.009)		(0.054)		(0.002)	
Unemployed	0.026		0.971**		0.095**	
	(0.054)		(0.309)		(0.013)	
Education	−0.092**		−0.642**		−0.014**	
	(0.013)		(0.075)		(0.003)	
Female	−0.133**		−0.577**		−0.033**	
	(0.036)		(0.204)		(0.009)	
Age	0.002		0.029**		−0.001**	
	(0.001)		(0.006)		(0.0003)	
% Immigrants (2007)	−1.434		−10.017*		−0.208	
	(0.748)		(4.282)		(0.186)	
% Unemployed (2007)	−2.472**		−17.881**		0.244	
	(0.740)		(4.232)		(0.184)	
Urban	0.022		0.041		0.001	
Model	G-est. Stage 1	G-est. Stage 2	G-est. Stage 1	G-est. Stage 2	G-est. Stage 1	G-est. Stage 2
Current state FEs (N=16)	✓		✓		✓	
Observations	1,376	1,376	1,376	1,376	1,376	1,376
Adjusted R^2	0.117	0.025	0.215	0.042	0.240	0.016

Note: Entries are coefficients of the effect of distance to closest camp on different outcomes, described in the column headers. Model 1, 3, and 5 correspond to the first stage of the sequential g-estimation (standard errors in parentheses), with contemporary covariates including current state fixed effects. Models 2, 4, and 6, represent the second stage in the sequential g-estimation (bootstrapped standard errors in parentheses). *p<0.05; **p<0.01

Table SM4.2: Effects of camp proximity on out-group intolerance, immigrant resentment, and support for far-right parties (EVS), accounting for systematic differences across Weimar states (interpolated from centroids of contemporary Kreise)

	Out-group Intolerance		Immigrant Resentment		Support Far-Right Parties	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to camp	-0.011^{**}	-0.010^{*}	-0.077^{**}	-0.051^{*}	-0.001	-0.002
(in 10kms)	(0.003)	(0.004)	(0.020)	(0.023)	(0.001)	(0.001)
% Jews (1925)	-1.876	-0.580	8.532	16.023	-0.077	0.183
	(1.192)	(1.494)	(7.037)	(9.978)	(0.268)	(0.474)
% Unemployed (1933)	2.344^{**}	3.128^{**}	0.143	11.169	0.107	0.682^{**}
	(0.719)	(0.915)	(4.243)	(5.799)	(0.162)	(0.237)
Population (1925)	-0.037^{**}	0.005	-0.300^{**}	-0.064	-0.003	-0.001
	(0.012)	(0.014)	(0.071)	(0.099)	(0.003)	(0.003)
Nazi party share (1933)	-0.466^{*}	-0.418	-2.796^{*}	-5.806^{**}	-0.021	-0.082
	(0.196)	(0.217)	(1.156)	(1.490)	(0.044)	(0.056)
Model	OLS	G-est. Stage 2	OLS	G-est. Stage 2	OLS	G-est. Stage 2
Weimar state FEs (N = 15)	✓	✓	✓	✓	✓	✓
Contemporary variables		✓		✓		✓
Observations	2,055	1,375	2,055	1,375	2,055	1,375
Adjusted R^2	0.045	0.057	0.075	0.096	0.005	0.034

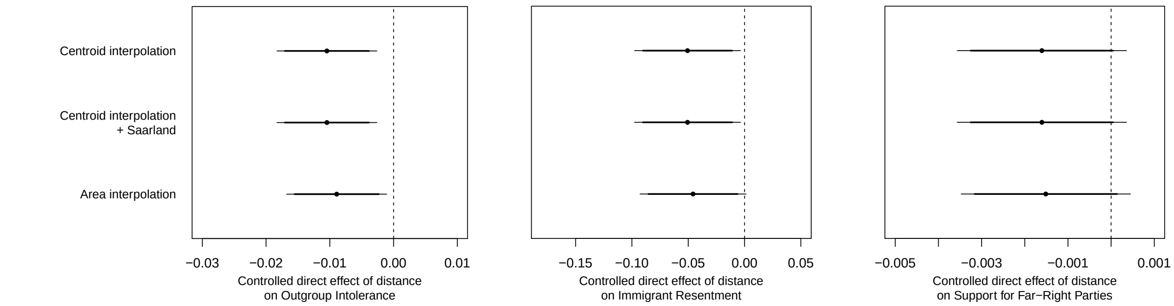
Note: Entries are estimates of the effect of distance to closest camp on the different outcomes, described in column headers. Models 1, 3, and 5 account exclusively for interwar covariates (standard errors in parentheses). Models 2, 4, and 6 are the 2nd stage of the sequential g-estimator to also account for contemporary predictors (bootstrapped standard errors in parentheses). All models (and both stages of the g-estimator) include Weimar state fixed effects (interpolated based on the centroids of contemporary Kreise). $^{*}p<0.05$; $^{**}p<0.01$

Table SM4.3: The controlled direct effect of camp proximity on out-group intolerance, immigrant resentment, and support for far-right parties (EVS), with current state fixed effects in first stage of g-estimator and accounting for systematic differences across Weimar states (interpolated from centroids of contemporary Kreise)

	Outgroup Intolerance		Immigrant Resentment		Support Far-Right Parties	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to camp	0.005	−0.018**	−0.021	−0.097**	−0.0005	−0.003*
	(0.005)	(0.005)	(0.030)	(0.027)	(0.001)	(0.001)
% Jews (1925)	10.570	−1.690	26.777	9.171	1.462	0.002
	(6.733)	(1.486)	(38.371)	(9.552)	(1.671)	(0.461)
% Unemployed (1933)	2.329*	3.142**	2.808	7.614	0.315	0.604*
	(1.088)	(0.971)	(6.200)	(6.010)	(0.270)	(0.247)
Population (1925)	−0.024	−0.007	−0.144	−0.173	−0.003	−0.003
	(0.017)	(0.016)	(0.098)	(0.111)	(0.004)	(0.004)
Nazi party share (1933)	−0.650*	−0.249	−4.861**	−4.632**	−0.166*	−0.055
	(0.288)	(0.229)	(1.642)	(1.636)	(0.071)	(0.060)
Contemporary covariates						
Conservatism	0.049**		0.297**		0.039**	
	(0.009)		(0.054)		(0.002)	
Unemployed	0.026		0.933**		0.094**	
	(0.054)		(0.309)		(0.013)	
Education	−0.092**		−0.647**		−0.014**	
	(0.013)		(0.075)		(0.003)	
Female	−0.133**		−0.559**		−0.034**	
	(0.036)		(0.204)		(0.009)	
Age	0.001		0.028**		−0.001**	
	(0.001)		(0.006)		(0.0003)	
% Immigrants (2007)	−1.773*		−12.225**		−0.296	
	(0.790)		(4.503)		(0.196)	
% Unemployed (2007)	−2.599**		−19.576**		0.085	
	(0.768)		(4.380)		(0.191)	
Urban	0.019		0.073		0.003	
	(0.013)		(0.073)		(0.003)	
Model	G-est. Stage 1	G-est. Stage 2	G-est. Stage 1	G-est. Stage 2	G-est. Stage 1	G-est. Stage 2
Current state FEs (N=16)	✓		✓		✓	
Weimar state FEs (N=15)	✓	✓	✓	✓	✓	✓
Observations	1,375	1,375	1,375	1,375	1,375	1,375
Adjusted R^2	0.118	0.039	0.221	0.080	0.245	0.019

Note: Entries are coefficients of the effect of distance to closest camp on different outcomes, described in the column headers. Model 1, 3, and 5 correspond to the first stage of the sequential g-estimation (standard errors in parentheses), with contemporary covariates including current state fixed effects. Models 2, 4, and 6, represent the second stage in the sequential g-estimation (bootstrapped standard errors in parentheses). All models also include Weimar state fixed effects (interpolated based on the centroids of contemporary Gemeinden) in *both* stages. $^{*}p<0.05$; $^{**}p<0.01$

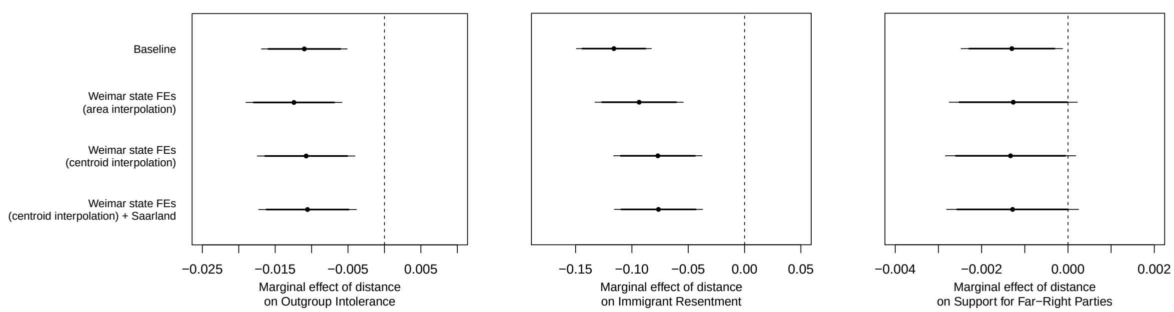
Figure SM4.1: The controlled direct effect of camp proximity on outgroup intolerance, immigrant resentment, and support for far-right parties (EVS), with alternative methods to interpolate Weimar states from contemporary Kreise



(a) Out-group Intolerance **(b)** Immigrant Resentment **(c)** Far-Right Support

Note: Plots depict estimates and 95/90% confidence intervals from the sequential g-estimator for the controlled direct effects of distance to camps on contemporary attitudes (described in each panel). Each estimate corresponds to a different model based on alternative methods to match contemporary districts with Weimar-era states. *Centroid interpolation* estimates come from Models 2, 4, and 6 in Table SM4.2; *Centroid interpolation + Saarland* estimates from Table DA.3 in the Dataverse Appendix; and *Area interpolation* estimates from Table DA.4.

Figure SM4.2: Effects of camp proximity on out-group intolerance, immigrant resentment, and support for far-right parties (EVS), accounting for state-level heterogeneity



(a) Out-group Intolerance **(b)** Immigrant Resentment **(c)** Far-Right Support

Note: Plots depict estimates and 95/90% confidence intervals from OLS models for the effects of distance to camps on contemporary attitudes (described in each panel label). Each estimate corresponds to a different model based on alternative methods to match contemporary districts with Weimar-era states. *Baseline* estimates correspond to the results reported in Table 2 (Models 1, 3, and 5) in HPT. *Centroid interpolation* estimates come from Models 1, 3, and 5 in Table SM4.2; *Centroid interpolation + Saarland* estimates from Table DA.5 in the Dataverse Appendix, and *Area interpolation* estimates from Table DA.6.

SM5: Prussian provinces: discussion and additional analyses

PGZ's original critique (2020) includes an analysis with Weimar-era state fixed effects. Although the authors use a map from 1925, the results reported are substantively similar to those we report here and inconsistent with PGZ's own argument. They then further split the analyses, treating Prussian internal provinces as separate states, which renders the effects of camp proximity unreliable. In the most recent version of the critique, PGZ only report the models with a combination of Weimar states and Prussian provinces. In this section, we take a closer look at these analyses and point out that the decision to split Prussia into provinces appears arbitrary and atheoretical. We then proceed to show that when we include province-level fixed effects in our electoral analysis, the main findings also remain unchanged – something PGZ did not report.

1. *Arbitrary decision to include Prussian provinces:* PGZ's main argument for the inclusion of Prussian provinces in the analyses with Weimar-era fixed effects is that provinces are the historical antecedents of contemporary states. This argument makes sense if the goal is to identify regions that match the geography of contemporary states as closely as possible. However, the goal of accounting for regional differences in a model is to absorb heterogeneity that (a) is of theoretical interest and (b) results from the socioeconomic and political variation in the regions *before* exposure to treatment. Trying to identify pre-treatment areas based on the shapes of post-treatment areas goes against this idea. Moreover, the decision to include fixed effects for Weimar-era Germany's *states* and Prussia's *provinces* in the same model specification seems arbitrary. Such a setup is equivalent to including state fixed effects for some US states, and county (or congressional district) fixed effects for others. Although California might be bigger and more heterogeneous than Rhode Island, we are not aware that this is a common empirical approach.[3]

2. *Atheoretical approach:* PGZ claim that within-Prussia heterogeneity necessitates the inclusion of province fixed effects. As an example, in the original critique the authors discuss a failed Reichsreform which planned to divide Prussia into several sub-states. We actually believe this example makes an argument *against* the inclusion of province fixed effects. The reform failed because Prussia did not want to lose the influence it had over the federal government (as the largest state). It was the opposition of the

[3]Another comparison could be the inclusion of fixed effects for the nine English regions (i.e., East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, and Yorkshire and the Humber) alongside fixed effects for Scotland, Wales, and Northern Ireland in studies focusing on the UK. We are not aware that this is a common approach either.

Prussian state as a whole – not of specific provinces – that blocked this reform (Schulz 1963). As such, the failed reform is an example of Prussian unity, not division. Holborn (1956: 335) explicitly describes that in "Prussia itself [...] no strong signs could be found that the provinces wished to become states." In fact, the Weimar constitution included a provision for the possible secession of individual provinces through plebiscite. As far as we can tell, this instrument was used only once, in Upper Silesia, where over 90% voted to remain part of Prussia in 1922 (Hertz-Eichenrode 1969; Schattkowsky 1994). Another attempt to schedule a plebiscite in Hanover in 1924 failed because there was not enough interest among voters (Funk 2010; Heimann 2011). This is in line with other accounts that emphasize the relevance and strength of Prussia as a whole (Orlow 1991). In other words, these historical accounts provide little evidence to support strong concerns regarding Prussia's own inherent regional heterogeneity.[4] We believe this discussion highlights once again the importance of having well articulated theoretical arguments to motivate one's modeling choices. For instance, if one is interested in capturing differences in school curricula that are determined at the state level, state-level fixed effects should be included. If instead one is interested in capturing administrative differences that vary at the substate level, then the analysis should include fixed effects at the level of the administrative region below the state.[5] Importantly, choosing one level in one state and another in the remaining states is a decision that seems difficult to motivate theoretically.

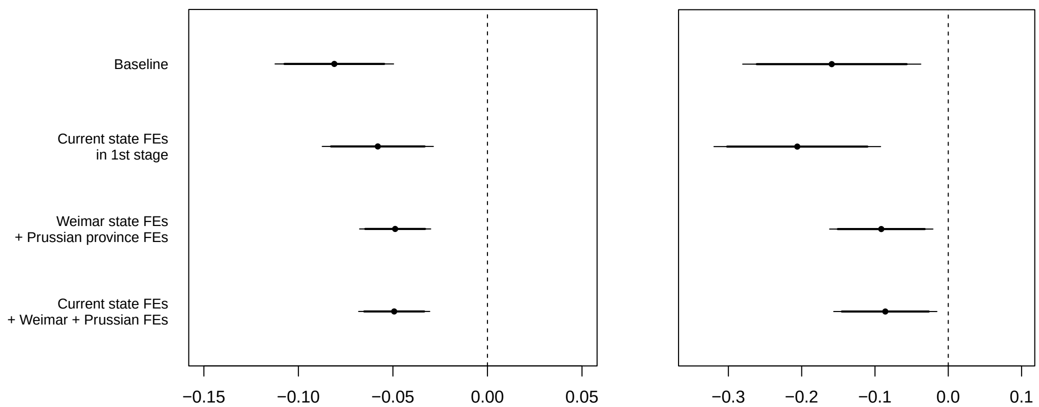
3. *Prussian provinces and electoral data:* Despite our concerns regarding PGZ's arbitrary and atheoretical inclusion of province fixed effects, we decided to replicate our electoral analysis while accounting for Prussia's provinces. More specifically, we follow the analysis in the main text and replicate Table 4 in HPT while including (1) contemporary state-level fixed effects in the first stage of the g-estimator, (2) Weimar-era state-level and Prussian province-level fixed effects in both stages of the g-estimator, and (3) both contemporary states *and* Weimar-era states and Prussian provinces. The results in Figure SM5.1 show that our main conclusions are robust to this hybrid approach of using fixed effects at the province level within Prussia and at the state level for the rest of the country. Across the different specifications, we see that the effect of distance is

[4] This does not mean that all of Prussia was always perfectly united. For example, there was a limited and ultimately unsuccessful independence movement in the Rhine Province that led to the declaration of a short-lived "Rhenish Republic" in the mid-1920s (Epstein 1967).

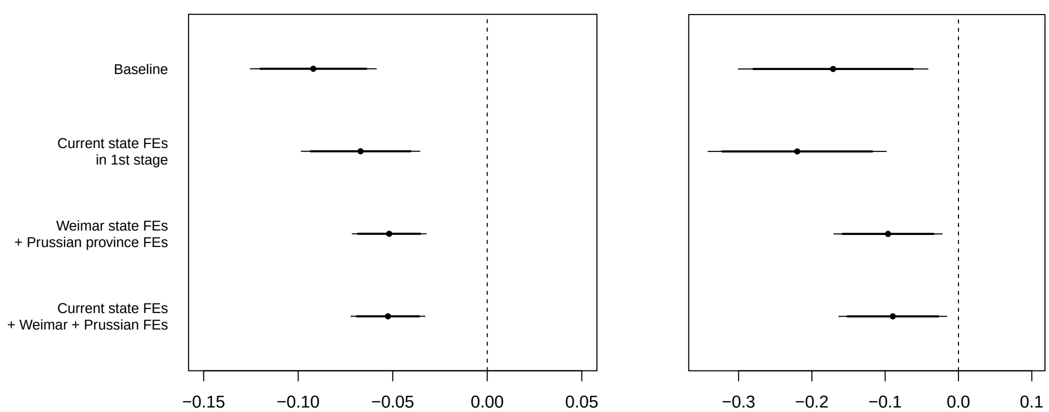
[5] In the case of Weimar Germany, choosing a level below the state is more challenging because of the differences in administrative setups across states. However, all states were ultimately divided up into a combination of Ämter, Kreise, and Regierungsbezirke, which would allow for the inclusion of fixed effects at that level.

always negative and statistically reliable at conventional levels. PGZ neglect to report these results.

Figure SM5.1: The controlled direct effect of camp proximity on support for radical right parties in 2017, accounting for state-level and Prussian province-level heterogeneity



(a) AfD, full sample **(b)** AfD, 70km radius



(c) AfD+NPD, full sample **(d)** AfD+NPD, 70km radius

Note: Plots depict estimates and 95/90% confidence intervals from the sequential g-estimator for the controlled direct effects of distance to camps on support for radical right parties in 2017 (described in each panel label). Each estimate corresponds to a different model specification, described on the y-axis. *Baseline* estimates come from Table 4 in HPT, *Current state FEs in 1st stage* estimates come from Table SM3.1, *Weimar state FEs + Prussian province FEs* estimates come from Table DA.7 in the Dataverse Appendix, and *Current state FEs + Weimar + Prussian province FEs* estimates come from Table DA.8.

SM6: Noise simulations

PGZ's critique is motivated in part by growing concerns about the effects of spatial correlation in the historical legacies literature (Kelly 2019). Although state-level heterogeneity – when properly incorporated in the analyses – does not explain the findings in HPT, spatial correlation may still play a role. To directly assess the robustness of findings to spatial correlation, Kelly suggests reestimating the main models with spatially-correlated noise.

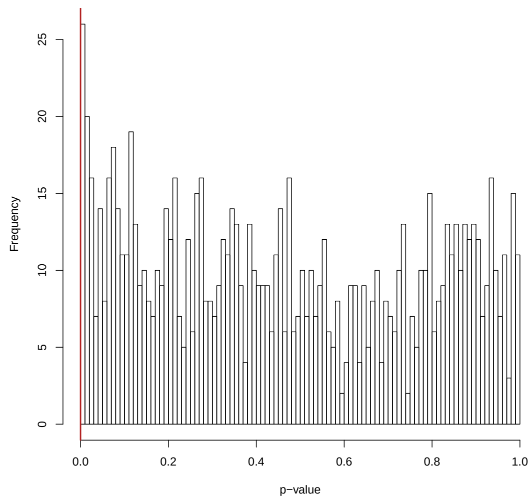
In this section, we conduct noise simulations to investigate the extent to which the main findings in HPT might be explained by spatial noise. Following Fouka and Voth (2023), we replace the geographic variable – *Distance* – with spatially correlated noise. The procedure consists of running 1,000 replications of the main models, each time with a different vector of spatially correlated random noise replacing *Distance*. Following Kelly, this spatial noise is drawn from a multivariate normal distribution using a variance-covariance matrix based on the Matern function. For the Matern function, we set the variance and shape parameters to be 1. For the crucially important correlation range, we follow Kelly's recommendations for the analysis of German data and present our results for a correlation range of 3 degrees and a correlation range of 5 degrees.

For each observation, we use this setup to draw 1,000 iterations of spatial noise. We then run 1,000 regressions replicating our g-estimation models, where the spatial noise replaces the *Distance* variable.[6] For every regression, we store the p-value of the spatial noise variable. The distribution of these p-values is then plotted in Figures SM6.1-SM6.4, along with a red vertical line illustrating the p-values from the original regressions. Across the different datasets and outcome variables, spatial noise very rarely outperforms our *Distance* variable in terms of explanatory power.

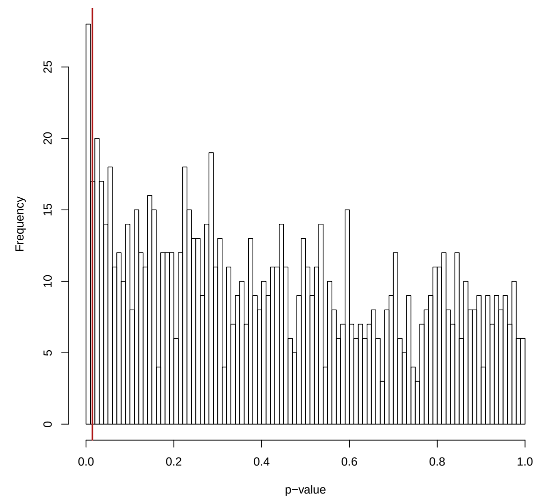
Table SM6.1 summarizes the main results from this simulation exercise. More specifically, it shows the amount of times that spatial noise had more explanatory power than HPT's original treatment variable across the 1,000 simulations. The results suggest that spatially-correlated noise rarely outperforms the *Distance* variable. As a reference point, in most of Kelly's replications of papers based on European data, the explanatory power of spatial noise outperformed the original predictor in 20-50% of all cases. None of HPT's models approaches these numbers. Spatial noise outperformed camp proximity less than 5% of the time in 11 of the 14 sets of simulations performed. The only exceptions are the electoral data models within 70km radii and a 5 degrees correlation range, where noise outperformed the predictor 11% of the time. The analyses suggest that camp proximity captures something more meaningful than mere spatial noise.

[6] Given the computationally intensive nature of this method, we report regular (i.e., non-bootstrapped) standard errors for these models.

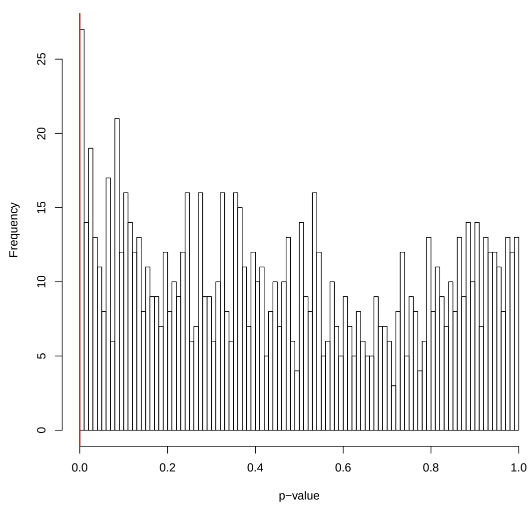
Figure SM6.1: The controlled direct effect of spatial noise on support for radical right parties in 2017 (correlation range of 3 degrees)



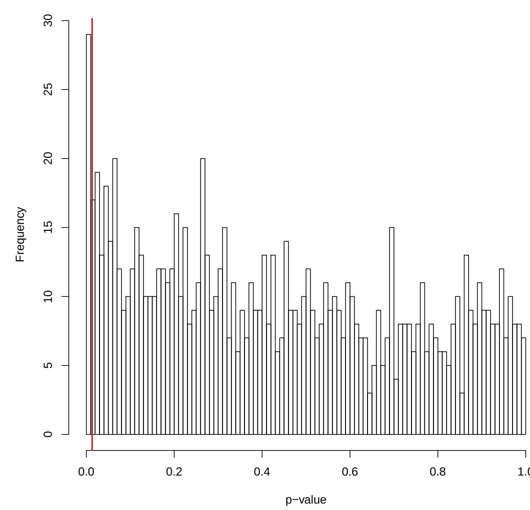
(a) AfD, full sample



(b) AfD, 70km radius



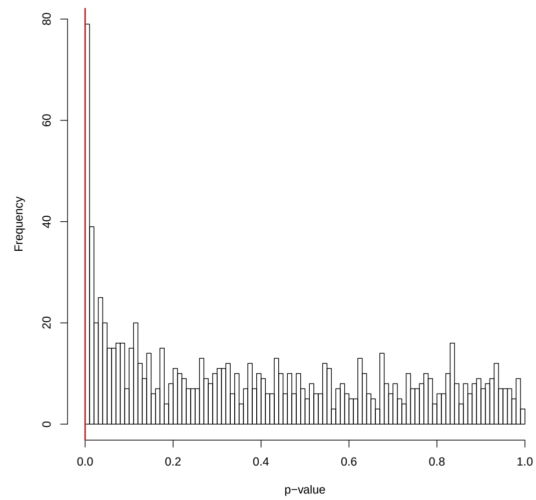
(c) AfD+NPD, full sample



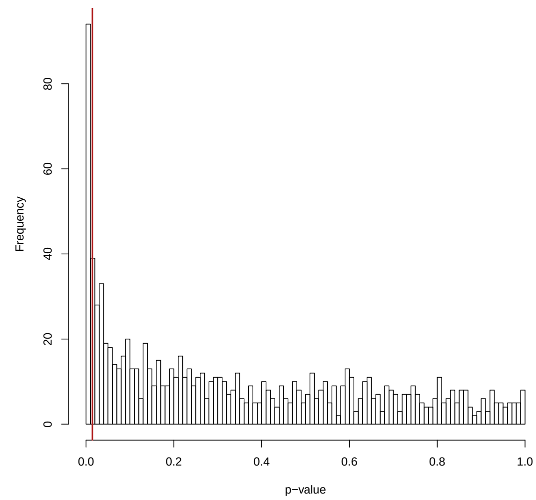
(d) AfD+NPD, 70km radius

Note: Plots show the distribution of p-values resulting from 1,000 simulations of the sequential g-estimator for the controlled direct effects of distance to camps on support for radical right parties in 2017 (described in each panel label). In each iteration, *Distance* has been replaced by simulated spatially correlated noise according to the Matern function, with a variance and shape of 1 and a correlation range of 3 degrees, following Kelly (2019) and Fouka and Voth (2023). The red vertical line indicates the p-value from the original model specification. The explanatory power of spatial noise is higher than that of our *Distance* variable 0.0% (panel a), 3.8% (panel b), 0.0% (panel c), and 3.4% (panel d) of the time respectively.

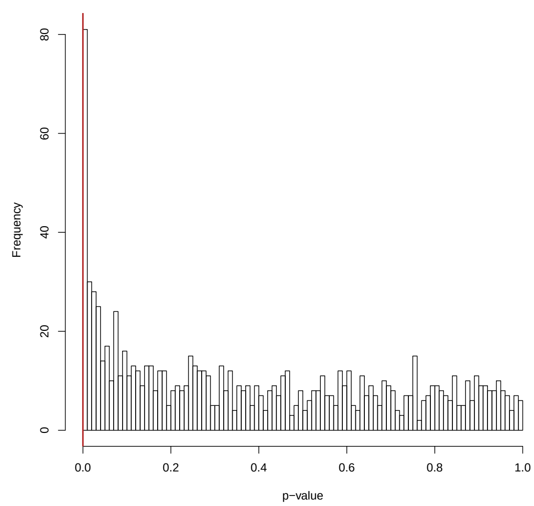
Figure SM6.2: The controlled direct effect of spatial noise on support for radical right parties in 2017 (correlation range of 5 degrees)



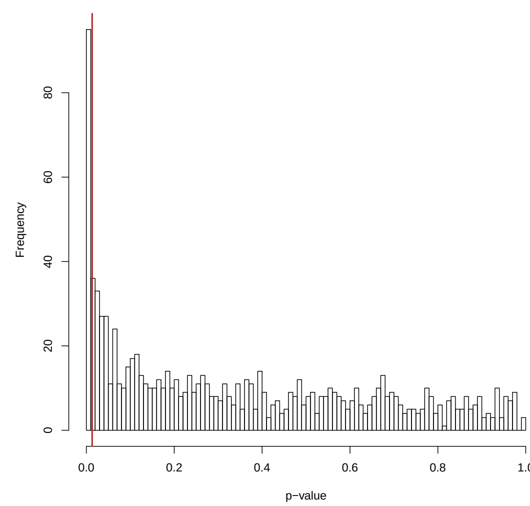
(a) AfD, full sample



(b) AfD, 70km radius



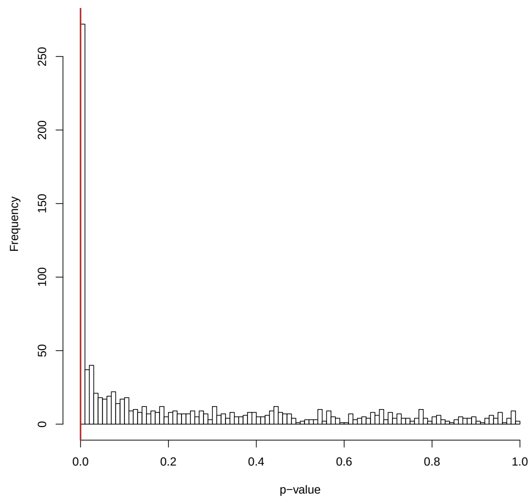
(c) AfD+NPD, full sample



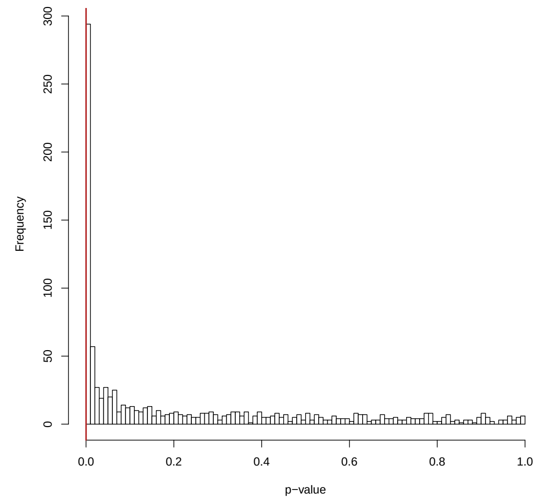
(d) AfD+NPD, 70km radius

Note: Plots show the distribution of p-values resulting from 1,000 simulations of the sequential g-estimator for the controlled direct effects of distance to camps on support for radical right parties in 2017 (described in each panel label). In each iteration, *Distance* has been replaced by simulated spatially correlated noise according to the Matern function, with a variance and shape of 1 and a correlation range of 5 degrees, following Kelly (2019) and Fouka and Voth (2023). The red vertical line indicates the p-value from the original model specification. The explanatory power of spatial noise is higher than that of our *Distance* variable 0.0% (panel a), 11.2% (panel b), 0.0% (panel c), and 11.1% (panel d) of the time respectively.

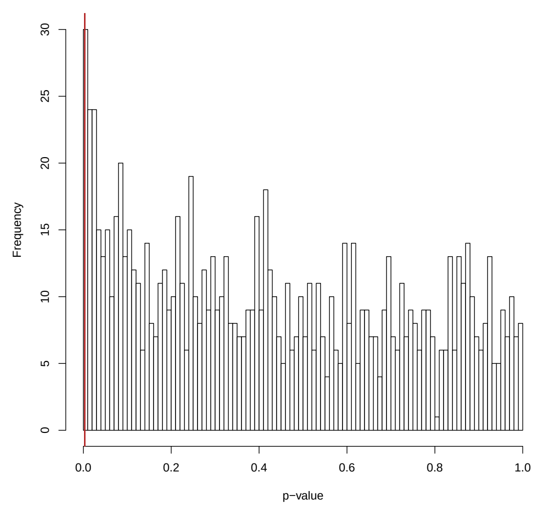
Figure SM6.3: The controlled direct effect of spatial noise on outgroup intolerance, immigrant resentment, and support for far-right parties (correlation range of 3 degrees)



(a) Outgroup Intolerance



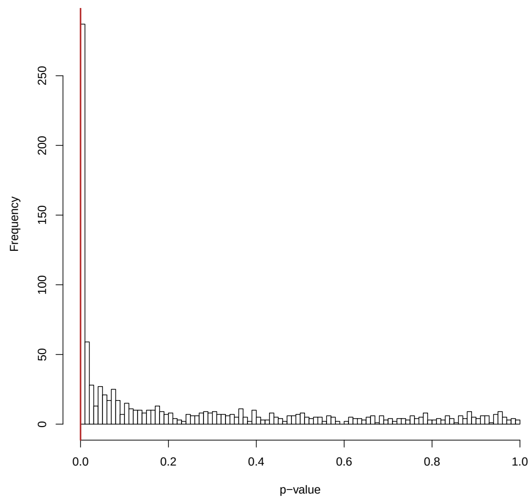
(b) Immigrant Resentment



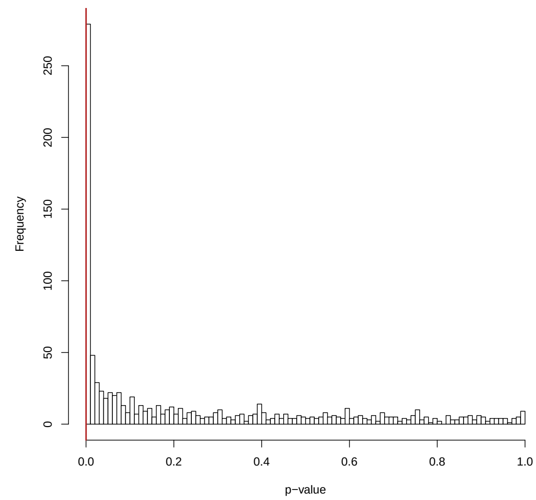
(c) Far-Right Support

Note: Plots show the distribution of p-values resulting from 1,000 simulations of the sequential g-estimator for the controlled direct effects of distance to camps on contemporary attitudes (described in each panel label). In each iteration, *Distance* has been replaced by simulated spatially correlated noise according to the Matern function, with a variance and shape of 1 and a correlation range of 3 degrees, following Kelly (2019) and Fouka and Voth (2023). The red vertical line indicates the p-value from the original model specification. The explanatory power of spatial noise is higher than that of our *Distance* variable 5.1% (panel a), 0.4% (panel b), and 1.1% (panel c) of the time respectively.

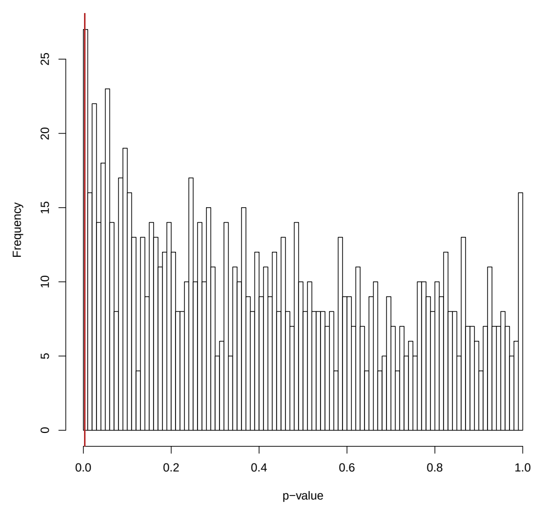
Figure SM6.4: The controlled direct effect of spatial noise on outgroup intolerance, immigrant resentment, and support for far-right parties (correlation range of 5 degrees)



(a) Outgroup Intolerance



(b) Immigrant Resentment



(c) Far-Right Support

Note: Plots show the distribution of p-values resulting from 1,000 simulations of the sequential g-estimator for the controlled direct effects of distance to camps on contemporary attitudes (described in each panel label). In each iteration, *Distance* has been replaced by simulated spatially correlated noise according to the Matern function, with a variance and shape of 1 and a correlation range of 5 degrees, following Kelly (2019) and Fouka and Voth (2023). The red vertical line indicates the p-value from the original model specification. The explanatory power of spatial noise is higher than that of our *Distance* variable 4.5% (panel a), 0.9% (panel b), and 1.1% (panel c) of the time respectively.

Table SM6.1: Explanatory power of the effect of spatial noise on contemporary outcomes

	Correlation range	
	3 degrees	5 degrees
Electoral data		
AfD, full sample	0.0%	0.0%
AfD, 70km radius	3.8%	11.2%
AfD+NPD, full sample	0.0%	0.0%
AfD + NPD, 70km radius	3.4%	11.1%
EVS		
Outgroup intolerance	5.1%	4.5%
Immigrant resentment	0.4%	0.9%
Far-right support	1.1%	1.1%

Note: Entries indicate how often spatial noise outperforms the original distance variable in explaining the different contemporary outcomes.

SM7: References

Epstein, Klaus. 1967. "Adenauer and Rhenish Separatism." *The Review of Politics* 29: 536-545.

Fouka, Vasiliki, and Hans-Joachim Voth. 2023. "Collective Remembrance and Private Choice: German–Greek Conflict and Behavior in Times of Crisis." *American Political Science Review* 117: 851-870.

Funk, Albert. 2010. *Kleine Geschichte des Föderalismus: vom Fürstenbund zur Bundesrepublik.* Paderborn: Ferdinand Schöningh.

Heimann, Siegfried. 2011. *Der Preußische Landtag 1899-1947: eine politische Geschichte.* Berlin: Links.

Hertz-Eichenrode, Dieter. 1969. *Politik und Landwirtschaft in Ostpreußen 1919-1930: Untersuchung eines Strukturproblems in der Weimarer Republik.* Köln, Opladen: Westdeutscher Verlag.

Holborn, Hajo. 1956. "Prussia and the Weimar Republic." *Social Research* 23: 331-342.

Kelly, Morgan. 2019. "The Standard Errors of Persistence." June 3. Available at SSRN. DOI: `10.2139/ssrn.3398303`

Orlow, Dietrich. 1991. *Weimar Prussia, 1925-1933: The Illusion of Strength.* Pittsburgh: University of Pittsburgh Press.

Schattkowsky, Ralph. 1994. "Separatism in the Eastern provinces of the German Reich at the end of the First World War." *Journal of Contemporary History* 29: 305-324.

Schulz, Gerhard. 1963. *Zwischen Demokratie und Diktatur: Verfassungspolitik und Reichsreform in der Weimarer Republik. Band 1. Die Periode der Konsolidierung und der Revision des Bismarckschen Reichsaufbaus 1919-1930.* Berlin: de Gruyter.