Electoral Studies 89 (2024) 102789

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Oppression of Catholics in Prussia does not explain spatial differences in support for the radical right in Germany. A critique of Haffert (2022)☆

Kai Arzheimer *, Theresa Bernemann, Timo Sprang

Department of Political Science, Johannes Gutenberg University Mainz, Germany

ARTICLE INFO

Keywords:
Radical right
Germany
Voting behaviour
Political geography
Ecological fallacy
Historical legacies
Political Catholicism

ABSTRACT

A growing literature links contemporary far-right mobilization to the "legacies" of events in the distant past, but often, the effects are small, and their estimates appear to rely on problematic assumptions. We re-analyse Haffert's (2022) study, a key example of this strand of research. Haffert claims that historical political oppression of Catholics in Prussia moderates support for the radical right AfD party among Catholics in contemporary Germany. While the argument itself has intellectual merit, we identify some severe limitations in the empirical strategy. Retesting the study's cross-level interaction hypothesis using more suitable multi-level data and a more appropriate statistical model, we find a modest overall difference in AfD support between formerly Prussian and non-Prussian territories. However, this difference is unrelated to individual Catholic religion or to the contextual presence of Catholics. This contradicts the oppression hypothesis. Our study thus provides another counterpoint to the claim that historical events have strong and long-lasting effects on contemporary support for the radical right. We conclude that simpler explanations for variations in radical right support should be exhausted before resorting to history.

1. Introduction

An emerging literature uses quantitative methods to link regional historical events to contemporary political developments and, in particular, to far-right mobilization. Often, these studies focus on Germany, whose traumatic history makes accounts of such persistence particularly poignant. While intellectually stimulating, there are three problems with this strand of research: the focus on long-term historical explanations can distract from simpler explanations based on proximate causes, the effects in question are usually very small, and their identification relies on strong and sometimes problematic methodological assumptions (see, for example, Pepinsky et al., 's 2023 critique of Homola et al., 2020 and Guinnane and Hofmann's 2022 rejoinder to Voigtländer and Voth, 2012 and Satyanath et al., 2017).

Perhaps the best example of these "legacy studies" is Haffert (2022), who claims that political oppression of Catholics in the 19th century is responsible for spatial disparities in support for the radical right AfD party observed in 2017 in the western states of Germany. This contribution is particularly interesting for two reasons: First, the religious cleavage was central for the development of many European party systems, but its relevance for the radical right party family is rarely discussed. Second, it explicitly addresses the thorny question of which vertical and horizontal transmission mechanisms could possibly link the historical experience of a group to the behaviour of its current members.

However, Haffert's study also suffers from some of the limitations outlined above. While the article's historical-theoretical premise is innovative and the reconstruction of the purported causal link is more elaborate than in most comparable contributions, the empirical evidence provided to support it is ultimately not convincing. More specifically, at the heart of the argument sits a *cross-level interaction hypothesis* about the voting behaviour of individual Catholic voters in different parts of Germany that is primarily tested with *macro data*. As we will show, this approach does not adequately reflect the structure of the theoretical argument and carries a serious risk of committing an ecological fallacy.

Accordingly, results from a replication with more adequate data and a statistical model that closely reflects the theoretical argument does not support the main claims made in Haffert (2022). By using a large-scale geo-referenced election survey to directly estimate the cross-level interaction effect posited by Haffert, we find a modest overall

☆ Kai Arzheimer received funding (grant number AR 369/4-1) from the German Science Foundation (DFG) for the collection of data used in this research note.
* Corresponding author.
E-mail addresses: arzheimer@politik.uni-mainz.de (K. Arzheimer), bernemann@politik.uni-mainz.de (T. Bernemann), sprang@politik.uni-mainz.de (T. Sprang).

https://doi.org/10.1016/j.electstud.2024.102789
Received 5 December 2023; Received in revised form 3 April 2024; Accepted 15 April 2024
Available online 27 April 2024

difference in support for the radical right 'Alternative for Germany' (AfD) between formerly Prussian and other areas that is however unrelated to both individual religion and to the contextual presence of Catholics. In other words, there is no empirical support for the notion that contemporary Catholics' behaviour towards the radical right can be linked to their forebears' political oppression 150 years ago.

2. The historical-oppression argument in a nutshell

The structure of the argument developed in Haffert (2022) closely follows the template of an ideal sociological explanation championed by Coleman.

1. Through political events at the *macro level* — the Congress of Vienna (1814–15) and the founding of the Second German Empire (1870) — a number of predominantly Catholic territories (most importantly the Rhineland) came under the control of the Protestant Prussian state. Catholics in these territories faced legal and political oppression by the authorities. In other predominantly Catholic areas that remained under the control of Catholic governments (e.g., in Bavaria), there was no comparable persecution of Catholics (Haffert, 2022, 599).
2. At the *meso level*, Catholics within Prussia responded by forming a dense network of political and other associations that was dominated by lay persons. Outside Prussia, Catholic organizations were generally weaker, less political, and dominated by the clergy (Haffert, 2022, 599–600).
3. Still at the meso level, this politicised network or "milieu" persisted throughout the Second Empire, the Weimar Republic, the "Third Reich" and even the first decades of the Federal Republic. In somewhat diminished form, it prevails to the present day and retains political norms that are inimical to radical right mobilization (*vertical* transmission). These norms were and still are *horizontally* transmitted to Catholics (Haffert, 2022, 600).
4. This implies a testable cross-level interaction: because these norms in turn affect *micro level* (voting) behaviour, Catholics in formerly Prussian areas should be less likely to vote for the AfD than Catholics in other parts of West Germany. Put differently, the *Prussian legacy of an area, via a history of oppression and the politicisation of the Catholic milieu, moderates the strength of the micro level relationship between denomination and voting behaviour*. In the following, we will focus on this last element of the presumed causal chain, because it provides the crucial link between past events and current political behaviour.

Before we turn to a discussion of the empirical strategy employed in Haffert's study in the next section, we would like to flag up two points in the argument that are not fully elaborated. First, it is not entirely clear if the cross-level interaction is confined to *predominantly* Catholic areas. The article refers several times to regions that are "majority Catholic" (e.g., Haffert, 2022, 596), but does not explicitly state this as a scope condition or introduce a three-way interaction in its core model. Second, Haffert (2022, 600) suggests in passing that the cross-level interaction might also affect lapsed Catholics or non-Catholics. Again, there is no further discussion and no attempt to model such a spill-over. In our view, an effect that is not confined to Catholics would also defeat the core theoretical argument, which highlights the transmission of norms within a closed Catholic milieu. We will address both points in the discussion of our own model.

3. The problems with the empirical strategy employed in Haffert (2022)

To test the macro-meso link, Haffert first regresses indicators of Catholic mobilization on a measure of historical anti-Catholic oppression. He finds correlations that drop from 0.55 in 1913 to 0.36 in the post-war period (Haffert, 2022, 607).

For the contextually moderated link between denomination and voting for the radical right, he then regresses the AfD's 2017 vote share in 8370 German incorporated municipalities (*Gemeinden*) on a set of state fixed effects, a host of controls (e.g., unemployment and immigration), the share of Catholics, a dummy for formerly Prussian territories, and an interaction between the latter two variables.[1] This interaction is negative. More specifically, the share of Catholics is *negatively* associated with the AfD result in formerly Prussian municipalities, whereas this association is *positive* in non-Prussian territories (Haffert, 2022, 603–4).

The paper takes this finding — complemented by a second series of aggregate analyses that zoom in on a former border region and an analysis of micro data from the 2017 election study that we will discuss below — as evidence that the cross-level interaction hypothesis holds, i.e. *that Catholics in some areas of western Germany are less inclined to* support *the radical right than in others, because Catholics in these parts faced oppression in the* 19th *century*. However, the data and main model suffer from serious limitations and hence cannot support this claim.

Our main criticism concerns the fact that the core model in Haffert (2022) effectively amounts to a naïve form of ecological regression, an attempt to infer micro behaviour from aggregated (meso level) data, whereas the central hypothesis posits an *interaction between micro and meso level variables* and hence requires a full multilevel model that *combines* micro data and contextual information.

Ecological inference, even if confined to bivariate micro-level relationships, is riddled with problems that ultimately stem from the loss of micro level information through aggregation. It relies on strong assumptions and specific techniques to be valid (King et al., 2004). While the main consequence — the danger of committing an ecological fallacy — is mentioned briefly in the article, the problem itself is hardly discussed in the text nor addressed in the main analysis.[2]

The study's choice of areal units exacerbates the problem. Aggregate correlations are more likely to reflect micro level correlations when units are uniformly small and homogeneous with respect to the categories of the purported independent variable, and when the micro level relationship is constant across units (an assumption violated by differences between predominantly Catholic and other regions). In this regard, German municipalities are not ideal even for purely descriptive aggregate analyses. Because the regulation of local government is a matter of state law, incorporated municipalities vary wildly across states in numbers as well as in terms of average size and political power. Two states alone (Rhineland-Palatinate and Schleswig-Holstein) collectively account for 3403 (40%) of all West German municipalities, although

[1] In a second series of models, replacing the dummy for Prussian territories with the more specific but spatially coarser measure for oppression events leads to broadly similar results, so we will focus on the former.

[2] In a bid to hedge against the danger of committing an ecological fallacy, Haffert presents (in Appendix C) an additional set of regressions estimated only on those ~2500 (30% of the full sample) municipalities where Catholics make up at least two thirds of the population. He finds that the Prussia effect is somewhat stronger in this sub sample and takes this as additional evidence. We cannot agree with this interpretation, because limiting the sample exacerbates some of the problems outlined in the main text below without overcoming the fundamental limitations of aggregate analysis. First, in the revised sample, Rhineland-Palatinate now accounts for 34% of the cases (an increase from 27%), and an additional 49% are located in Bavaria. Conversely, the Catholic population of North Rhine-Westphalia, which is larger than that of Bavaria in absolute terms, is represented by just 3% of the cases. Even worse, important and populous centres of political Catholicism in North Rhine-Westphalia such as Bonn, Cologne, or Essen are excluded because they do not meet the 2/3 criterion. However, with a median number of 963 voters per municipality, there is still a considerable danger of aggregation bias, and even in this highly selective sub sample, most municipalities are not nearly homogeneous: the share of Catholics at the ninth decile is 88%, so most municipalities are not very informative with respect to the political behaviour of their Catholic denizens.

Table 1
Regressions on AfD vote.

	(1) Municipality	(2) County	(3) Electoral District
AfD vote			
Male	0.405***	0.410***	0.407***
	(0.0695)	(0.0638)	(0.0639)
Education: high	−0.465***	−0.470***	−0.468***
	(0.0853)	(0.0815)	(0.0807)
30-60	0.153	0.152	0.151
	(0.130)	(0.139)	(0.141)
60+	−0.301*	−0.301*	−0.303
	(0.145)	(0.153)	(0.155)
manual/service	0.426***	0.430***	0.432***
	(0.0807)	(0.0817)	(0.0786)
Catholic	−0.257	−0.342	−0.275
	(0.420)	(0.428)	(0.450)
Protestant	−0.682*	−0.857**	−0.766*
	(0.344)	(0.320)	(0.340)
Prussian	−0.327	−0.335	−0.307
	(0.220)	(0.238)	(0.243)
Catholic × Prussian	0.438	0.689	0.513
	(0.572)	(0.582)	(0.616)
Protestant × Prussian	0.225	0.355	0.165
	(0.432)	(0.448)	(0.450)
% Catholics	0.00429	0.00507	0.00384
	(0.00694)	(0.00718)	(0.00762)
Catholic × % Catholics	0.00283	0.00496	0.00262
	(0.0100)	(0.0100)	(0.0103)
Protestant × % Catholics	0.0132	0.0184*	0.0152
	(0.0103)	(0.00927)	(0.00971)
Prussian × % Catholics	0.000648	0.000146	0.000816
	(0.00766)	(0.00766)	(0.00812)
Catholic × Prussian × % Catholics	−0.0160	−0.0222	−0.0164
	(0.0139)	(0.0143)	(0.0144)
Protestant × Prussian × % Catholics	−0.0115	−0.0154	−0.00906
	(0.0130)	(0.0129)	(0.0128)
% Unemployment	0.149**	0.133*	0.121*
	(0.0520)	(0.0574)	(0.0551)
% Marginal employment	−0.0484	−0.109*	−0.0812
	(0.0495)	(0.0529)	(0.0614)
% Regular employment	0.00345	−0.0184	−0.0188
	(0.0280)	(0.0283)	(0.0298)
Population (1,000s)	0.000141	0.0000878	0.000156
	(0.000168)	(0.000125)	(0.000192)
University town/city	−0.102	−0.0818	−0.102
	(0.112)	(0.0952)	(0.106)
% Foreigners	−0.00298	−0.00167	0.00693
	(0.0105)	(0.0122)	(0.0127)
% 65+	0.0366	0.0247	0.0504
	(0.0254)	(0.0339)	(0.0328)
% Female	−0.0780	−0.104	−0.162
	(0.0682)	(0.0681)	(0.0875)
Close to border	0.340	0.319	0.427
	(0.328)	(0.344)	(0.309)
Constant	0.890	3.841	5.845
	(3.825)	(3.923)	(4.740)
Var. Constant	3.33e-33	2.26e-33	1.45e-33
	(1.45e-32)	(6.91e-33)	(1.66e-33)
Observations	10,969	10,969	10,969
Clusters	699	274	214

Standard errors in parentheses.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

they are home to just 10% of the western population. Conversely, North Rhine-Westphalia with a population of about 18 million people (26%), has less than 400 incorporated municipalities (5%). The other states fall somewhere along this spectrum. That makes municipalities a problematic unit of analysis.

This is also evident in the vast differences in population sizes: while the biggest western municipalities (Cologne, Hamburg, Munich) boast populations that exceed one million, the smallest municipalities have populations in the low double figures. To address this problem, Haffert weighs his data with the number of voters, but this strategy leads to quite extreme weights. Smaller municipalities with fewer than 350 voters make up one quarter of the sample. They are weighed down by factors between 0.32 and 0.03, although the loss of information through aggregation is comparatively small here. Conversely, the quarter of municipalities that have 3400 voters or more are given weights that are at least ten times higher, namely in the range between 3.1 and 895 (sic!).

Worse still, most municipalities are heterogeneous with respect to the focal variable. Catholics make up only 35% of the population in the median municipality. Even at the upper quartile, their share is just 71%. The vast majority of municipalities is therefore not even informative regarding the simple relationship between being Catholic and voting for the AfD, let alone the moderating effect a contextual variable might have on this. Moreover, a proper test of the interaction hypothesis cannot be confined to the Catholic segment of the population – a point that we will elaborate in our discussion of our own model.

Aggregate correlations should therefore be treated with extreme caution. They reflect a mixture of (1) Catholics' individual behaviour, (2) contextual effects of Catholicism, and (3) aggregation bias, whose respective proportions are unknown. In line with these concerns, the core finding of a *positive macro level relationship between Catholicism and voting AfD in the non-Prussian territories* (Haffert, 2022, 604) does not reflect the *well-documented negative micro level relationship between church membership and radical right voting* in Germany and the wider West European context (Marcinkiewicz and Dassonneville, 2022).

Fortunately, it is not necessary to rely on ecological inference to test the cross-level hypothesis. In Haffert (2022, 610–11), there is a supplementary micro level model based on data from the 2017 GLES election survey. Regrettably, this analysis has some serious shortcomings, too. Because of the limited sample size, it does not include state effects and meso level control variables (not even the Catholic share of the population), making it prone to omitted variable bias. Moreover, historical oppression is measured at a different, considerably higher level of aggregation, and a truncated propensity to vote variable is used as a proxy for AfD voting, because too few AfD voters were sampled. Finally, the estimation is restricted to Catholic respondents. This means that any effects of living in formerly Prussian areas that affect non-Catholics and are therefore incompatible with the core theoretical mechanism are ruled out *a priori*. Hence, the validity of this analysis is severely limited.

As an alternative, we propose a fully specified multilevel model that better reflects the core theoretical arguments and allows for a direct test of the crucial cross-level interaction.

4. A multilevel model of religion and AfD voting in 2017

On our behalf, Infratest Dimap administered a CAWI survey three months before the 2017 German election.[3] For this survey, a large (n = 25,479) quota sample of German nationals, stratified by gender, age, education, and region, was drawn from a very large (n > 130,000) access panel whose members had been recruited offline.

Because of Germany's strict rules on data protection, geo-information about respondents' addresses had to be coarsened in low population density areas to prevent any risk of identification. However, we could still match 11,251 respondents to 702 of the municipalities in Haffert's dataset, amounting to 16 respondents per municipality on average. This should give us adequate power to detect cross-level interaction effects.

Although we have no respondents from many of the small and very small municipalities in Germany, this is does not pose a serious problem, because relatively few people live in self-governing hamlets, villages, or small towns, and because we are ultimately interested in the behaviour

[3] Replication data and scripts are available from the first author's dataverse at https://doi.org/10.7910/DVN/507J3S or directly as supplementary data from the web version of this article.

of voters, not municipalities.[4] As the definition and size of municipalities varies so much across states, the impact of excluding small municipalities on the representation of rural voters is relatively limited,[5] and we have little reason to assume that it biases our results.[6]

Because the core argument posits a cross-level moderating effect of historical oppression on AfD voting, we specify a logistic multilevel model that mirrors this argument as closely as possible (see Table 1). Our binary dependent variable is the intention to vote for the AfD in the 2017 election. The share of self-declared AfD voters was 7.7%, somewhat below the eventual result in the western states.

Since multilevel models do not automatically take care of heteroscedasticity (Hazlett and Wainstein, 2022), we employ cluster-robust standard errors throughout. For the most direct comparison with Haffert's core model, we treat voters as nested within municipalities (column 1 in Table 1). To address the concerns regarding the suitability of municipalities, we also repeat this analysis using two higher-level geographies: counties (column 2) have comparable political structures and powers throughout the nation but still vary considerably in terms of size and population. Electoral districts (column 3) have no political powers of their own but are required by law to have populations of a broadly similar size. By extending the analysis to the county and electoral district level,[7] we maximise the chance of finding evidence for a cross-level interaction and safeguard against the modifiable areal unit problem (MAUP).

To capture historical legacies, we use the dummy variable for formerly Prussian territories, which has stronger and more consistent effects across all models than the alternative measure and is available in higher spatial resolution (Haffert, 2022, 601). We measure religion at the micro level by including a dummy variable for Catholic respondents. An interaction term between both variables provides a direct test for the cross-level interaction hypothesis.

To better reflect Germany's religious heterogeneity, and because the core argument is essentially about the *difference* between Catholics and other groups, we include a second dummy variable for members of the federation of mainstream Protestant churches (EKD) that we also interact with the Prussia variable, while any other respondents form the reference group.[8]

Because Haffert's contribution (sometimes) distinguishes between predominantly and non-predominantly Catholic areas within and outside former Prussia and to mirror his aggregate model, we also include the share of Catholics and interact this variable with the Prussia variable, too. We also include interactions between the share of Catholics and the individual dummy variables that measure religion, because the behaviour of the members of a social group is often conditional on majority/minority status.[9] Finally, as a safeguard against potential bias (Brambor et al., 2006), we include the two three-way interactions between the Prussia indicator, the Catholic/Protestant dummies, and the share of Catholics.[10]

At the contextual level, we include fixed state effects (omitted from the main table to conserve space but documented in the appendix, Table A.1) and all other controls from the main model in Haffert (2022, 603). At the micro level, we follow Haffert's supplementary micro model in including controls for age, gender and education. We replace his control for household income (not included in our data) with a dummy for manual and routine service occupations.[11]

Table 1 presents the estimates for the three variants of the model.[12] Measuring the contextual variables at higher levels of aggregation yields results broadly similar to those for measurements at the municipal level and thus does not alter the general conclusions drawn from the models: The unemployment rate has a strong positive effect, while all other controls at the context level have no significant effects (except for share of marginal employment on the county level).

The effects for the control variables at the micro level closely resemble previous findings in the literature. Being male and being employed in a manual or routine service job has a strong positive effect on the likelihood of an AfD vote, whereas higher levels of formal education are associated with a comparable negative effect. The average marginal effect for each of the variables amounts to a difference of about three percentage points (e.g., 10% for respondents in the manual/routine service sector vs 6.8% for all others). The (less pronounced) effect of higher age is also in line with previous research, but is only significant when voters are treated as clustered in municipalities or counties.

Because of the interactions and the non-linearity of the model, we plot predictive margins to assess the interplay of religion and Prussian legacies (see Fig. 1). The three panels correspond to scenarios where the

[4] Roughly 4700 cases (some of which would have been embargoed) are dropped from the models because of missing values. About 4300 cases, or just under 20% of the full western sample, have full information but need to be excluded purely because of data protection rules. In the appendix, we present an additional robustness test which shows that using these cases in the models for the higher-level geometries does not substantially change the results.

[5] Employing the classification developed by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR), which is defined at the county level and thus can be applied to all respondents, we find that roughly 80 per cent of all persons in the full sample are located in urban or suburban counties, while the remaining 20 per cent live in rural areas. Excluding respondents from the smallest municipalities and with missing values reduces the share of rural voters to 10 per cent.

[6] Following Haffert, our models include a number of contextual controls that should pick up structural effects of rurality (e.g. presence of higher education institutions, migration, and location near border). We also control for the most important individual predictors (age, gender, education) associated with demographic differences between urban and rural areas. Finally, we also re-ran model 2 (which does not require exact geolocation) on the full sample (see the appendix for the associated supplemental tables and figures). The results are virtually identical to those reported here. We are therefore confident that our results do not suffer from selection bias.

[7] For the higher-level geographies, we calculate all contextual variables as the population-weighted average of the municipal-level data provided in the replication archive for Haffert (2022).

[8] The models presented in Haffert (2022) do not include this second interaction and hence effectively constrain the Prussia effect on non-Catholics to zero. Therefore, they do not provide a balanced test for the historical-oppression argument.

[9] We refrain from including the share of Protestants in the model, because there is a moderate-to-strong negative correlation with the share of Catholics.

[10] Heisig and Schaeffer (2019) suggest including random slopes for the lower-level variables in a cross-level interaction. However, due to limited random variation between municipalities after introducing the random intercept, fixed state effects, and the municipality-level controls, we could not achieve convergence when estimating random slopes for the Protestant and Catholic dummies. Omitting random slopes *increases the likelihood* of detecting evidence favouring the legacies hypothesis, as the recommendation's aim is to make tests for cross-level interactions *more conservative* (Heisig and Schaeffer, 2019, 267).

[11] Household income is also prone to item non-response. In Haffert's micro analysis, more than 20% of the respondents are excluded because they had missing values for income.

[12] Multi-level models can be presented in a stepwise fashion, adding first the micro-level variables, then the contextual levels, and finally any (cross-level) interactions or other complications. While this practice is generally useful, we opted for only presenting the full model for three reasons. First, to safeguard against the MAUP, Table 1 already shows three variants of the full model. Second, our theoretical and empirical focus is precisely on the cross-level interactions between individual religion and living in formerly Prussian territories, whereas the other variables (whose effects are well-documented in the literature) chiefly act as controls. Third, these interactions need to be assessed graphically because of the non-linearities involved, and coefficients are not comparable across specifications with and without interactions. Readers who are interested in estimates for simpler, intermediate models can easily reproduce these making use of the replication materials.

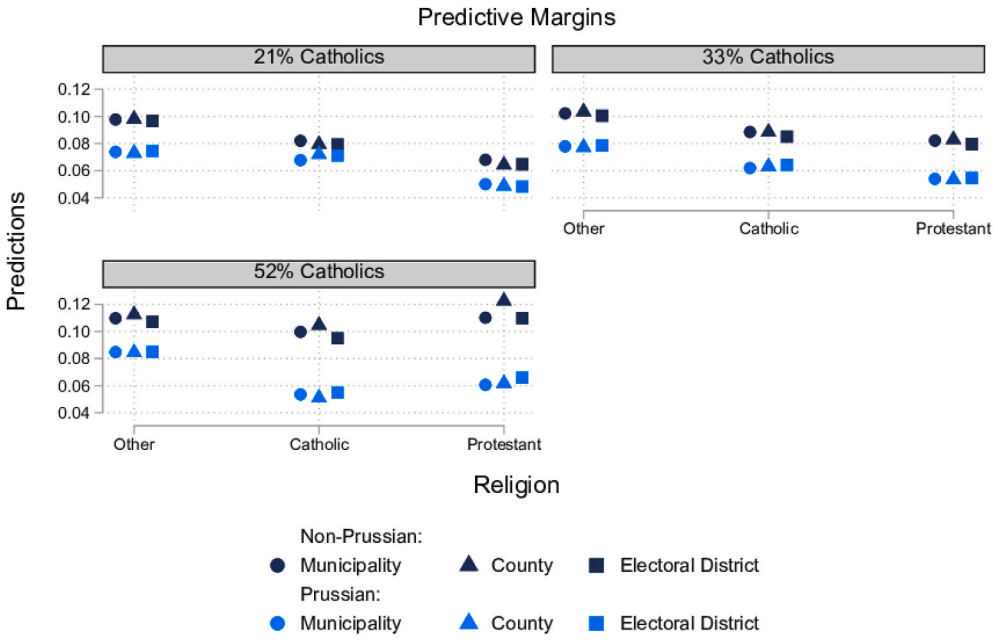

Fig. 1. Predictive marginal effect for voting AfD.

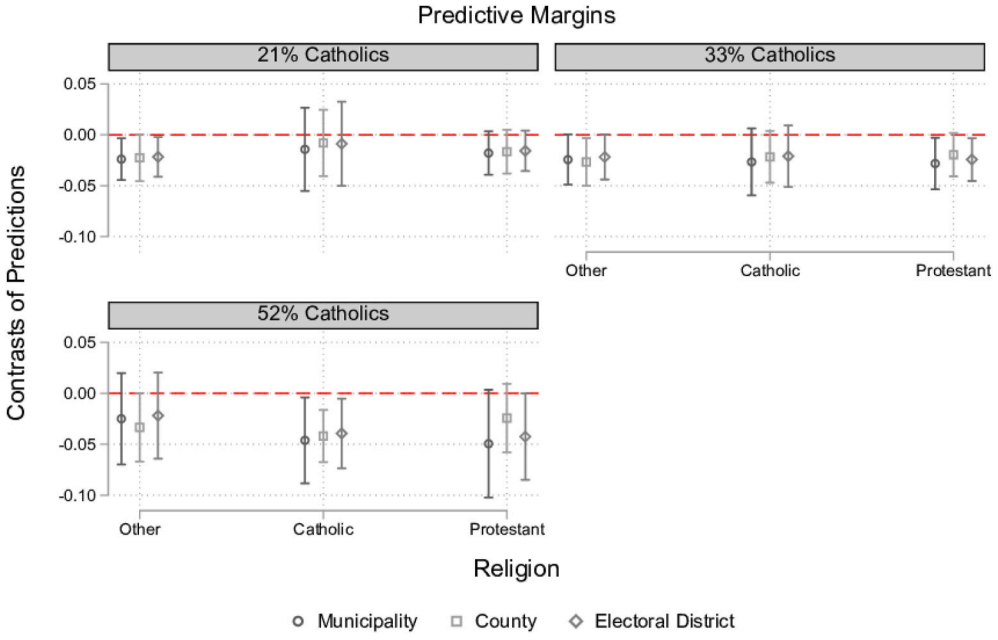

Fig. 2. Average marginal effects for voting AfD.

share of Catholics is at the lower quartile (21%), the median (33%), and the upper quartile (52%), respectively. Within each panel, we distinguish between (1) Catholics, Protestants, and other respondents, who are (2) living in formerly Prussian/non-Prussian territories.

A first result is that overall, differences between denominational groups amount to 2–3 percentage points at most, even though we do not control for any attitudes. Religion in itself is not a very strong predictor of radical right voting.

Turning to the cross-level interaction, the estimated levels of AfD support amongst Catholics are indeed somewhat lower in former Prussian territories than in non-Prussian across all three scenarios. However, as Fig. 2 reveals, this difference between Catholics in Prussian and non-Prussian areas is *only* significant within municipalities/counties/electoral districts where the Catholic share of the population is very high. In the other scenarios, the confidence interval includes the null value. Hence, even when the cross-level interaction is tested on more

appropriate contextual levels, the proposed effect is only visible in some rather exceptional[13] cases.

Moreover, this effect, which varies only marginally with the Catholics' population share, is by no means confined to Catholics. Across all three scenarios, the size of the gap is virtually identical for Protestants and "other" respondents and becomes significant for Protestants under the 33% scenario (and the 52% scenario in the county level model) and for "others" under the 21% scenario. This is incompatible with both the core mechanism and the secondary idea of a large Catholic milieu so influential that it even affects the behaviour of non-Catholics.

A Wald test confirms that one cannot reject the null hypothesis of all interactions being jointly zero. If anything, our results therefore show a modest, yet at 2.2 percentage points potentially politically relevant, effect of living in formerly Prussian territories that is not confined to any religious group and works independent of the presence of Catholics.

5. Discussion

Haffert (2022) is the strongest of several recent articles that seek to link contemporary radical right mobilization to the persistent effects of past events (see e.g. Homola et al., 2020; Voigtländer/Voth, 2012). It specifies plausible mechanisms for the vertical and horizontal transmission of electoral norms (even if there are some ambiguities) and closely follows the template of an ideal multi-level sociological explanation.

But unfortunately, even in this contribution the empirical strategy does not do justice to the structure of the theoretical argument and can therefore not provide valid estimates for the political behaviour of Catholics and other religious groups in contemporary Germany. Using a better suited, well-powered dataset and a fully specified multilevel model, we find evidence of a modest difference in AfD support between formerly Prussian and non-Prussian territories. This effect is, however, not confined to Catholics and hence incompatible with the historical-oppression argument.

This leads to the question of what might explain the differences between formerly Prussian and other territories. One possible explanation lies in the Prussian reform movement that *preceded* the suppression events described in Haffert (2022) by some sixty years. Following its humiliating defeat in the Napoleonic wars, Prussia embarked on an ambitious program to modernise not just its army but also its whole state apparatus. Within a decade, Prussia built a relatively modern education system, ended serfdom, introduced economic reforms and gave full citizen rights to its Jewish population. It is at least conceivable that a legacy of effective institutions and liberalism from above still curbs the demand for radical right politics.

Events that took place about 45 years *after* the "Kulturkampf" could also be responsible for the differences observed today: During the Weimar period, Prussia was a haven of (relative) democratic stability. As a federal state, it managed to democratise its security apparatus. Prussia was even able to integrate the very structures of political Catholicism described in Haffert (2022) into its ruling coalition (Schulze and Dwyer, 2001).

Both (very rudimentary) historical explanations are compatible with the small but significant contemporary difference in AfD support that we observe between formerly Prussian territories and other regions in western Germany, and so may be many others. In our view, this illustrates a more general point. Radical right voting is, in a sense, overdetermined, with several plausible mechanisms leading to the same result (see also Arzheimer and Bernemann 2023). Amongst these, historical accounts, especially when they imply long arcs, exert a certain intellectual fascination. But such explanations also rely on strong assumptions, are often impossible to disentangle from alternative and equally plausible narratives, and usually explain, as in this case, relatively little variance. We therefore believe that scholars of the radical right should first exhaust the simpler and more proximate causes of right-wing support before turning to history, if at all.

[13] For reasons of data confidentiality, we have to remove respondents from small and very small municipalities (see note 4). This does hardly affect the municipalities' median share of Catholics (33% vs 35%), but reduces its variation because extremely high or low shares of Catholics are almost exclusively observed in small towns and villages. Once more, this illustrates that because of the somewhat arbitrary definition of municipalities, their properties do not represent the properties of the west German population very well. While 75% of the *municipalities* have populations where the share of Catholics is lower than 52%, 75% of our *respondents* live in municipalities where the share of Catholics in the population is 41% or less. Therefore, the 52% scenario is rather extreme, as it reflects the surroundings of just 12% of our respondents.

CRediT authorship contribution statement

Kai Arzheimer: Writing – review & editing, Writing – original draft, Resources, Methodology, Funding acquisition, Conceptualization. **Theresa Bernemann:** Writing – review & editing, Writing – original draft, Conceptualization. **Timo Sprang:** Writing – review & editing, Visualization, Software.

Declaration of competing interest

There are no other interests to declare.

Data availability

Replication data and scripts are available from the first author's dataverse at https://doi.org/10.7910/DVN/507J3S or directly as supplementary data from the web version of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.electstud.2024.102789.

References

Arzheimer, K., Bernemann, T., 2023. 'Place' does matter for populist radical right sentiment, but how? Evidence from Germany. Eur. Polit. Sci. Rev. 1–20. https://doi.org/10.1017/S1755773923000279. Published online 2023.

Brambor, T., William, R.C., Golder, M., 2006. Understanding interaction models. Improving empirical analyses. Polit. Anal. 14 (1), 63–82.

Guinnane, T.W., Hoffman, P., 2022. Persistence and historical evidence: the example of the rise of the Nazi party. In: CESifo Working Paper No. 10095. https://doi.org/10.2139/ssrn.4286968.

Haffert, L., 2022. The long-term effects of oppression: Prussia, Political Catholicism, and the Alternative für Deutschland. Am. Polit. Sci. Rev. 116 (2), 595–614. https://doi.org/10.1017/s0003055421001040.

Hazlett, C., Wainstein, L., 2022. Understanding, choosing, and unifying multilevel and fixed effect approaches. Polit. Anal. 30 (1), 46–65. https://doi.org/10.1017/pan.2020.41.

Heisig, J.P., Schaeffer, M., 2019. Why you should always include a random slope for the lower-level variable involved in a cross-level interaction. Eur. Sociol. J. 258–279. https://doi.org/10.1093/esr/jcy053.

Homola, J., Periera, M.M., Tavits, M., 2020. Legacies of the third Reich: concentration camps and out-group intolerance. Am. Polit. Sci. Rev. 114 (2), 573–590. https://doi.org/10.1017/S0003055419000832.

King, G., Rosen, O., Tanner, M.A. (Eds.), 2004. *Ecological Inference. New Methodological Strategies*. Analytical Methods for Social Research. Cambridge University Press, Cambridge.

Marcinkiewicz, K., Dassonneville, R., 2022. Do religious voters support populist radical right parties? Opposite effects in western and East-central Europe. Party Polit. 28 (3), 444–456. https://doi.org/10.1177/1354068820985187.

Pepinsky, T.B., Goodman, S.W., Ziller, C., 2023. Modeling spatial heterogeneity and historical persistence: Nazi concentration camps and contemporary intolerance. Am. Polit. Sci. Rev. 1–10. https://doi.org/10.1017/S0003055423000072.

Satyanath, S., Voigtländer, N., Voth, H.J., 2017. Bowling for fascism: social capital and the rise of the Nazi party. J. Polit. Econ. 125 (2), 478–526.
Schulze, H., Dwyer, P.G., 2001. Democratic Prussia in Weimar Germany, 1919-33. In: Dwyer, P.G. (Ed.), Modern Prussian History 1830-1947. Routledge, London, pp. 211–229.
Voigtländer, N., Voth, H.J., 2012. Persecution perpetuated: the medieval origins of anti-semitic violence in Nazi Germany. Q. J. Econ. 127 (3), 1339–1392. https://doi.org/10.1093/qje/qjs019.