Electoral Studies 96 (2025) 102957

Contents lists available at ScienceDirect

Electoral Studies

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

Research note

The East in wolf's clothing. Wolf attacks correlate with but do not cause far-right voting[☆]

Nico Sonntag

Institute of Sociology, Johannes Gutenberg University Mainz, Jakob-Welder-Weg 12, Mainz, 55128, RLP, Germany

ARTICLE INFO

Dataset link: https://osf.io/6uvnf/

Keywords:
Far-right parties
Voting
Germany
Wolf attacks
Green party
AfD
Parallel trends

ABSTRACT

The resurgence of wolves in Germany has sparked intense debate, particularly in rural areas where wolf attacks on livestock are frequent. Prior research has linked wolf attacks to a rise in support for the far-right AfD party and the corresponding decline in Green Party votes. This note challenges such conclusions, highlighting significant methodological issues in the original difference-in-differences design. The geographic clustering of wolf attacks in East Germany aligns with other pre-existing political divides, complicating causal attribution. The analysis reveals pre-treatment voting trends that predict the probability of wolf attacks. Splitting data by East and West Germany or including diverging regional trends in the regression models nullifies or reverses most previously reported effects. This study underscores the complexities of using panel data in spatially and temporally heterogeneous contexts.

1. Introduction

The return of the wolf (*Canis lupus*) is a contentious issue in rural German regions. The species had been nearly extinct in Germany since the mid-19th century. Over the last 20 years, it has made a remarkable comeback, which has been met with mixed reactions. Parts of the rural population — not just livestock owners — have responded with hostility or fear (Zscheischler and Friedrich, 2022). In a recent study, Clemm von Hohenberg and Hager (2022) examined the relationship between wolf attacks on livestock and regional voting patterns. They focused on the far-right AfD party (*Alternative für Deutschland*), which has frequently echoed public concerns about the wolf, as well as the environmentalist Green Party, which generally supports the conservation of wolves.[1] They found a rather strong relationship between wolf attacks on livestock and increases in far-right voting across federal, state-level, and municipal elections. At the same time, the Green Party's vote share decreases following wolf attacks. For example, in federal elections, the far-right AfD appears to gain on average up to 2.5 percentage points in municipalities that experience a wolf attack for the first time in the legislative period before the election. For state-level elections the effect seems even larger, "suggesting that the far-right AfD gains up to 5 percentage points after a wolf attack" (Clemm von Hohenberg and Hager, 2024a, 2).[2]

Clemm von Hohenberg and Hager's research question is intriguing because it moves beyond the preoccupation with socio-economic characteristics of individuals as well as the monothematic focus on the issue of immigration which have sometimes dominated the scientific discourse on the rise of the German far (or populist) right. Moreover,

☆ I thank Bernhard Clemm von Hohenberg and Anselm Hager for their openness in providing information and for a constructive discussion regarding my concerns. I thank Nicole Schwitter, Dave Balzer, Daniel Meyer and three anonymous reviewers for helpful feedback on earlier drafts of this research note.

E-mail address: sonntag@uni-mainz.de.

[1] There are some notable differences between the preregistered research design and final article which are never made transparent or explained: First, the authors preregistered only one hypothesis pertaining to the vote share of the Green party ("H1: Wolf attacks lower the vote share of the Green party") for which the analysis produces "inconsistent evidence" (Clemm von Hohenberg and Hager, 2022, 2). The main finding of the published article regarding "far-right voting" — that features prominently in the title ("Wolf attacks predict far-right voting") — is only mentioned as supplementary analysis that was to be tested "in an exploratory manner". In contrast, in the published version the authors explicitly state: "Our primary outcome of interest is the vote share of the far-right AfD, but we also analyze votes for the proenvironmental Green party". Results for the two other preregistered supplementary hypotheses ("H3: Wolf attacks increase (environmental) polarization" and "H4: Wolf attacks increase turnout") are not reported and not even mentioned in the final article. There might be good reasons to deviate from the preregistered analysis plan. However, in this case the differences should be disclosed and highlighted for otherwise the whole idea of a preregistered analysis is moot. Undisclosed deviations could actually decrease confidence in the findings and feed the suspicion that researcher degrees of freedom were exploited.

[2] The original analysis contained a coding error to which I alerted the authors. A correction has since been published (Clemm von Hohenberg and Hager, 2024a). I refer the corrected results throughout this research note.

https://doi.org/10.1016/j.electstud.2025.102957
Received 19 December 2024; Received in revised form 13 June 2025; Accepted 16 June 2025
Available online 7 July 2025

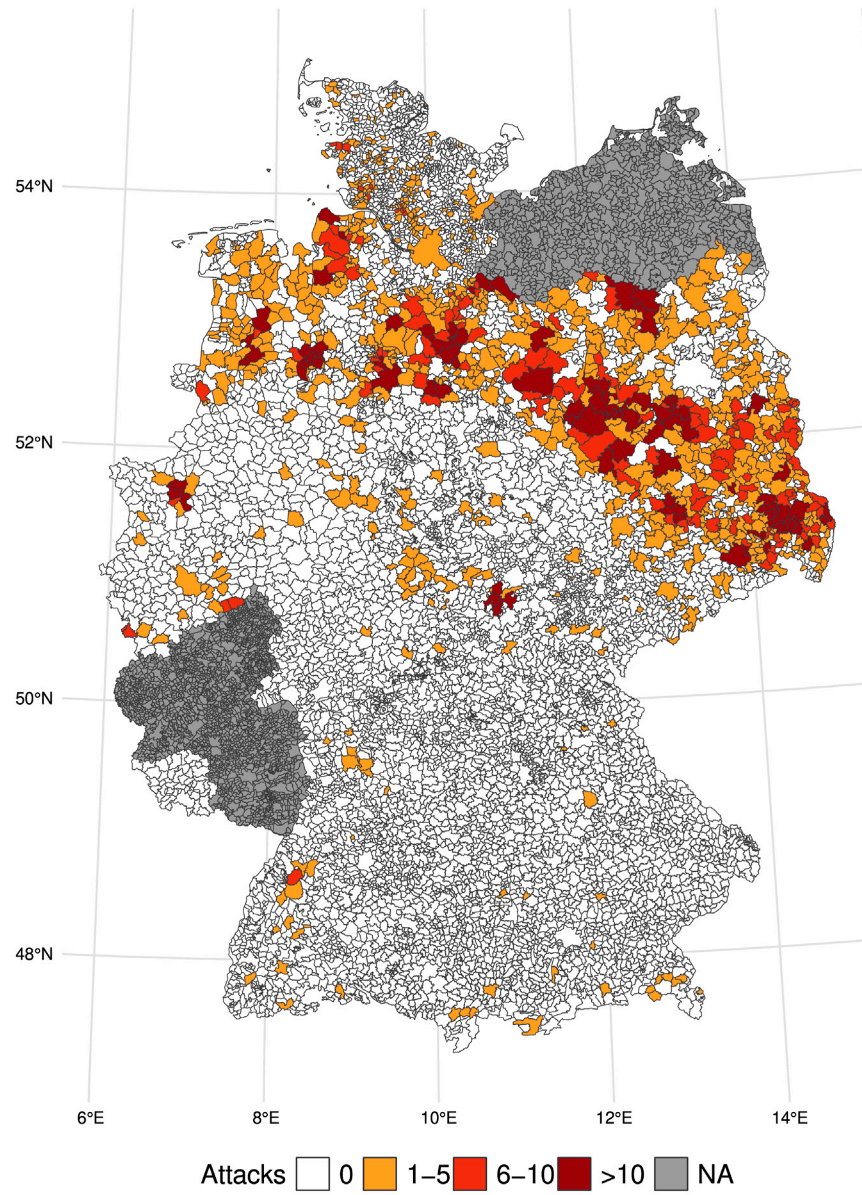

Fig. 1. Map of wolf attacks on livestock since 1998 at the municipality level (harmonized at the 2021 borders, $n = 10{,}976$; two states did not provide data). *Source:* Reproduced from Clemm von Hohenberg and Hager (2022) using the materials provided by the authors.

it connects a specific issue with local events, thus contributing to the explanation of regional voting patterns. Unfortunately, the results by Clemm von Hohenberg and Hager (2022) are uninformative with regard to their substantive research question. While it is true that "wolf attacks predict far-right voting" as stated in their subtitle, far-right voting also seems to predict wolf attacks, as I will show in more detail in Section 2. This is a problem since Clemm von Hohenberg and Hager (2022) do not merely report an interesting correlation. In the end, they are interested in whether wolf attacks *cause* a higher AfD vote share. They are careful to note that "one should be cautious in interpreting the findings in a causal manner" because their analysis "—although it absorbs a variety of likely confounders—still risks missing some sources of confounding" (Clemm von Hohenberg and Hager, 2022, 3). Nonetheless, one might hope, that the results provide at least strong *prima facie* evidence in favor of the hypothesized causal effects. However, their research design is vulnerable to a correlation between region-specific prior trends and treatment assignment. Accounting for divergent pre-treatment trends critically changes the central results, reversing the signs of the relationships or reducing their size to (almost) zero.

The underlying difference-in-differences design used in the analysis assumes parallel trends in the AfD vote share among all municipalities in the absence of wolf attacks. This assumption is strong and likely violated. As Fig. 1 reveals, wolf attacks cluster geographically in North-East Germany since the wolves that first recolonized Germany in the 90s originated from Poland (Andersen et al., 2015). 27% of all East German municipalities have recorded at least one wolf attack since 1998 compared to only 8% of West German municipalities. Meanwhile the AfD has been more successful in East Germany due to, for example, attitudinal and structural east–west differences completely unrelated to wolf attacks (Weisskircher, 2020; Arzheimer, 2023). It is likely that these factors did not just affect initial success (e.g. in the federal election of 2013) but also explain the stronger increase in vote share in East Germany compared to West Germany. Ignoring the possibility of diverging regional trends in AfD success over time very likely results in a mis-specified model. Unaccounted divergent trends in the two regions have caused problems in other research as well. For example, Schmidt-Catran and Spies (2016, 242) concluded based on German panel data "that native-born populations become more reluctant to support welfare programs when the proportion of foreigners at the regional level increases". However, Auspurg et al. (2019) demonstrated in a reanalysis that the association was driven by pre-existing trends among Eastern and Western respondents' support for welfare programs; a critique that was accepted by the authors of the original paper (Schmidt-Catran and Spies, 2019).

This research note examines *robustness reproducibility*, defined as the extent to which original study results remain robust under alternative, plausible analytical choices using the same data (Brodeur et al., 2024, 2–3). Proving the robustness of Clemm von Hohenberg and Hager's results necessarily involves the estimation of either separate models for East and West Germany or models including diverging trends for both regions. I focus on their main results. Their exploratory results are meant to make sense of the relationship between wolf attacks and voting patterns. However, since my reanalysis strongly suggests that this relationship was spurious, there is not much left to be explained by these additional analyses.

2. Far-right voting predicts wolf attacks

Clemm von Hohenberg and Hager's (2022) units of analysis are municipalities (n = 10,976) in the fourteen states for which data on wolf attacks are available (see Fig. 1). The dataset has a panel structure with multiple observations (election outcomes) for each municipality.[3] The observation period is 1990 to 2021 for the Green Party and 2013 to 2021 for the AfD. The dependent variable is, depending on the analysis, the vote share of either the AfD or the Greens in each respective municipality. Therefore, these are purely aggregate analyses. Since the units are not weighted by the number of votes cast, the target quantity is an average effect across municipalities, which is not equivalent to the total gain or loss in vote shares due to wolf attacks among the respective political electorate as a whole. The treatment is a dichotomous variable either indicating for each point in time whether a wolf attack has ever taken place in a municipality or indicating that an attack has happened in the preceding election period.[4]

The authors themselves test the crucial parallel trends assumption using an event study design (see also Section 3.2 below) but only for the Green Party (Clemm von Hohenberg and Hager, 2024b, 5). The AfD has only existed since 2013 and therefore "pre-treatment elections are too few to examine pre-trends meaningfully" (Clemm von Hohenberg and Hager, 2024b, 5). However, there are three observations per unit for federal elections — their "preferred outcome" (Clemm von Hohenberg and Hager, 2022, 2) — which allow for a comparison of pre-treatment trends from 2013 to 2017 for municipalities that remained untreated until at least the 2017–2021 election period.

Fig. 2 presents a simple visual check of the parallel trend assumption's credibility with regard to federal elections.[5] First, all municipalities that experienced wolf attacks before the 2017–2021 period were removed from the analysis sample. The remaining units are divided by their treatment status *in 2021*. Focusing on the AfD, even though there were on average no initial differences in the vote share in 2013, both groups were already diverging by the time of the 2017 federal election. This constitutes a non-parallel pre-treatment trend because, as already mentioned, Fig. 2 is based exclusively on municipalities that were not treated until after the 2017 election. Even though the difference between treatment and control group grows even stronger in 2021, we should be reluctant in attributing this to the causal effect of the treatment as there might be systematic unobserved differences between treatment and control group not accounted for by the removal of time-constant unit-specific errors. The pre-treatment trends for the Green Party diverge as well but to a lesser extent.[6] This is in line with Clemm von Hohenberg and Hager's own event studies which indicated diverging pre-treatment trends that "warrant caution" (Clemm von Hohenberg and Hager, 2022, 2).

The most obvious (observable but neglected) characteristic might be the municipalities' location. As mentioned in the introduction, wolf attacks are much more likely to happen in East Germany. Thus, it makes sense to look at the trends separately for both regions in order to check whether the parallel assumption seems more credible for regional subsamples.

Fig. 3 again plots the trends for AfD vote share by treatment status in 2021 but splits the sample into West and East German municipalities. For East Germany, the trends from 2013 to 2017 are not just parallel but the levels are actually similar. Levels are not strictly relevant for identification but the overall similarity lends some additional credibility to the assumption of counterfactually parallel trends. For West Germany, trends are still diverging prior to 2021, suggesting the possible relevance of other unobserved factors. As can be seen in Fig. 1, in addition to the East-West difference there is also a North-South gradient. The majority of all West German attacks happened in either Schleswig-Holstein or Niedersachsen. About 80% of all municipalities in West Germany where at least wolf attack has taken place until 2021 are located in these two North German states.

Fig. 4 repeats the analysis for Green Party vote shares. Again, not just trends but levels are more similar compared to the pooled sample. However, there are small but — given the effect size estimates for the Green Party — nonetheless potentially relevant pre-treatment differences. In West Germany the mean difference between treatment and control group is insignificant in both pre-treatment periods, while in East Germany the initially significant mean difference in 2013 vanishes by 2017.

To conclude, pre-treatment trends are potentially correlated with treatment assignment. In other words, changes in Green Party and far-right voting predict wolf attacks since vote share changes from 2013 to 2017 are correlated with attack probabilities in the next period (see Appendix). Accounting for East-West differences reduces the diverging trends but does not eliminate them. The AfD vote share among West German municipalities in particular still exhibits possibly confounding pre-treatment trends.

3. Reanalysis

Clemm von Hohenberg and Hager (2022, 3) specify a baseline (static) two-way fixed-effects (TWFE) model with an absorbing binary treatment

$$\text{vote}_{it} = \alpha_i + \beta_t + \delta \text{wolf}_{it} + \varepsilon_{it} \tag{1}$$

where vote_{it} is the AfD or the Green party vote share (in percent) in municipality i at election t; α_i are municipality fixed effects and β_t election fixed effects. wolf_{it} takes the value 1 if a municipality witnessed a wolf attack in period t before the election (remaining at

[3] Between 1990 and 2021 there have been 9 federal elections, 120 state elections, and 147 municipal elections in total.

[4] The treatment definition differs from the description in the preregistration materials. According to the latter, a municipality was to be coded as "treated" if it "has experienced a wolf attack in election-period, t" and the authors clarify "[t]he treatment only applies for one election period". The preregistration materials only mention the "absorbing" treatment used in the published article's main findings, i.e. the unit remains treated for the rest of the analysis period after it experienced its first wolf attack, as "an alternative model". However, both variants are highly correlated and results are unaffected by this choice.

[5] All analyses are based on the replication data provided by the authors (Clemm von Hohenberg, 2022). The data and Stata syntax files used in the reanalysis are available at https://osf.io/6uvnf/.

[6] The Green Party has been at the ballot for the full observation period (since 1990). However, since the majority of all wolf attacks took place since 2015 (Clemm von Hohenberg and Hager, 2022, 2), I focus on unequal trends before the 2021 election as well.

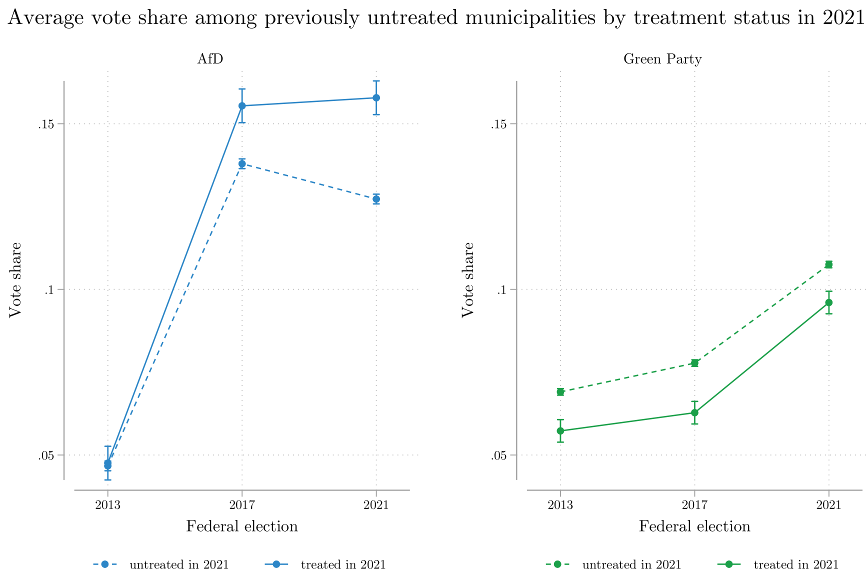

Fig. 2. Trends in party vote shares by treatment status in 2021. Municipalities that experienced wolf attacks before the 2017–2021 period were removed from the data. Data: Clemm von Hohenberg and Hager (2022); own calculation.

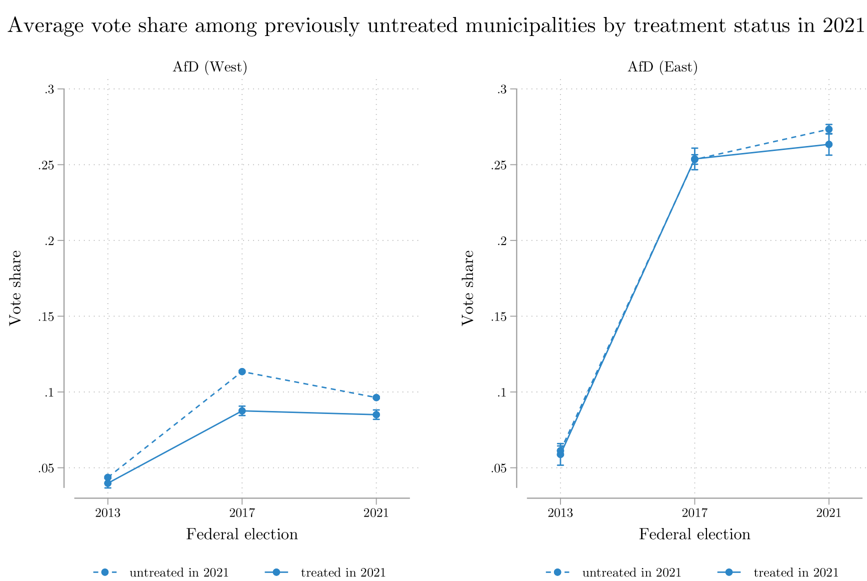

Fig. 3. Trends in AfD vote shares by region and treatment status in 2021. Municipalities that experienced wolf attacks before the 2017–2021 period were removed from the data. Data: Clemm von Hohenberg and Hager (2022); own calculation.

1 in the periods thereafter) and 0 otherwise. ε_{it} are the idiosyncratic error terms. The model parameters are estimated using the fixed-effects within-estimator and cluster-robust standard errors (clustered at the municipality level). Extended models by Clemm von Hohenberg and Hager (2022) include a number of time-varying control variables.[7]

The parameter δ identifies the average treatment effect on the treated (ATT) under the assumption of strict exogeneity of the idiosyncratic error term (Brüderl and Ludwig, 2014, 328–329). Strict exogeneity implies the parallel (or common) trends assumption, i.e. the trend in party vote shares of untreated units is used to estimate the counterfactual vote share of treated units. Because wolf attacks are geographically clustered in East Germany, where the AfD and Green vote share might have been on a different trend compared to West Germany independent of any wolf attacks, δ is potentially biased. Therefore, it is important to account for potentially different pre-treatment trends in the Eastern and Western parts.

In a first step, I keep the basic setup but re-estimate separate models for West and East Germany.[8] East Germany is defined as municipalities in the states of Brandenburg, Mecklenburg-Vorpommern, Saxony, Saxony-Anhalt, and Thuringia. Berlin is coded as belonging to West

[7] Clemm von Hohenberg and Hager (2022) estimate alternative model specifications using the panel matching estimator proposed by Imai et al. (2023). This estimator likewise relies on the parallel trends assumption and the general problem can be demonstrated by focusing on the more familiar DiD/TWFE specification.

[8] All specifications are estimated using the Stata ado `reghdfe` which implements an alternative (faster) algorithm for the multiway fixed-effects within-estimator (Correia, 2016, 2023).

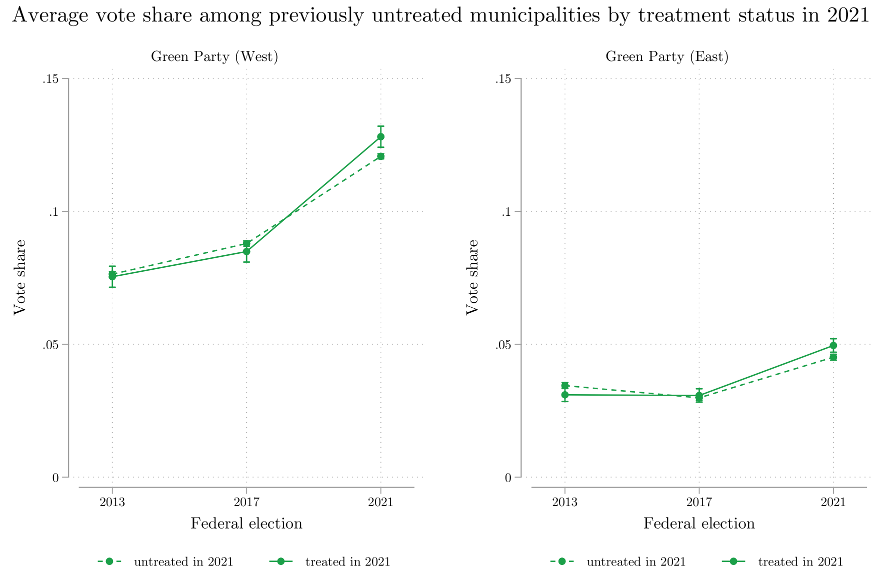

Fig. 4. Trends in Green Party vote shares by region and treatment status in 2021. Municipalities that experienced wolf attacks before the 2017–2021 period were removed from the data. Data: Clemm von Hohenberg and Hager (2022); own calculation.

Germany but results are robust to categorizing it as East Germany or to its removal.[9] In a second step, I pool the observations from both regions and model heterogeneous regional time trends in two different ways. In the first variant, election fixed effects vary by region r (dummy indicating East Germany), i.e. for each election there are two period-specific errors, one for West German and one for East German municipalities:

$$\text{vote}_{it} = \alpha_i + \beta_{0t} + \beta_{1t} r_i + \delta \text{wolf}_{it} + \varepsilon_{it}. \tag{2}$$

β_{0t} and β_{1t} — like α_i — are not explicitly modeled but "absorbed" by estimating a multiway fixed-effects model.

Moreover, as discussed in Section 2, in addition to East-West differences, there may also be a North-South gradient in wolf attacks, which is likely correlated with differing trends in AfD election outcomes. To account for this, I estimate a model that replaces the region-specific effects with state-by-period fixed effects:

$$\text{vote}_{it} = \alpha_i + \gamma_{s(i),t} + \delta\, \text{wolf}_{it} + \varepsilon_{it}. \tag{3}$$

Here, $s(i)$ denotes the state to which municipality i belongs. The term $\gamma_{s(i),t}$ therefore provides each state with its own election-specific intercept, i.e. one fixed effect for every state-period combination.

3.1. Main results

Fig. 5 shows the reproduced results for both parties from the basic TWFE specification (see Fig. 2 in Clemm von Hohenberg and Hager, 2024a, 2), separate models for both regions as well as results for the models described by Eqs. (2) and (3). Looking at the separate models, the positive association of wolf attacks with the AfD vote share disappears (West) or reverses (East). Moreover, the wolf attack coefficient turns positive for the Green Party vote share in both regions. Results for the two models that pool observations but attempt to account for diverging trends likewise contradict the hypothesized relationships. The signs are reversed in the specification with regional-specific election fixed effects. Hence, wolf attacks seem to *reduce far-right and increase proenvironmentalist Green voting*. However, these statistically significant results are unlikely to be true estimates of the ATT as well. Controlling for East-West differences removes a major source of non-parallel trends, yet other possible regional disparities remain (see Section 2).

State-by-period fixed effects flexibly absorb, for example, North-South variation, making this my preferred specification; with it, the null finding appears plausible. Critics may argue, however, that this approach "overdifferences" the data, that is, including too many fixed effects leaves too little variation (Cochrane, 2018). In the end, one can be agnostic whether any of the models is clearly preferable: What is clear is that the original results are not robust and likely stem from treatment assignment being confounded with regional time trends in party performance. Any attempt to mitigate diverging East-West trends yields estimates that contradict the original hypothesis.

Fig. 6 repeats the analyses for state elections. As before, simply splitting the sample into West and East nullifies the finding (West) or reverses (East) the sign for the AfD estimates. In both cases the absolute value is much smaller than the original estimate of almost 10 percentage points average vote share increase due to wolf attacks. Results from models accounting for diverging trends in the pooled sample via region-specific period fixed effects or state-by-period fixed effects are likewise not compatible with the originally proposed hypotheses. The "wolf effect" is zero or negative for the AfD in all specifications. The negative association between changes in Green Party vote share and wolf attacks is more robust except for the East German subsample. There remains a negative association in all other models, although its size is sometimes greatly reduced.

Lastly, results from local elections were never really in line with expectations. The coefficient for wolf attacks was positive for Green Party vote share in all specifications of the original analyses but turned insignificant at the 5% level due to covariate adjustment or matching (see Figure 2 in Clemm von Hohenberg and Hager, 2024a, 2). This general pattern holds throughout my reanalysis (Fig. 7). Without controls, coefficients are positive for both dependent variables even though their relative size changes with the exact specifications. The only exception is the model with state-by-period fixed effects. At the local level, wolf attacks appear to benefit either both parties or neither.

In summary, there is no consistent pattern across all estimates. For the AfD vote share, the relationship is either zero or negative for federal and state elections but positive for local elections. For the Green Party vote share, the relationship is positive for federal and local

[9] Due to its geographic location, it is equally reasonable to view Berlin as part of East Germany or remove it from the analysis altogether because it spans territory of both the former East and West Germany. Berlin's assignment to a region does not carry a lot of weight because it comprises a single municipality in the control group (no wolf attacks).

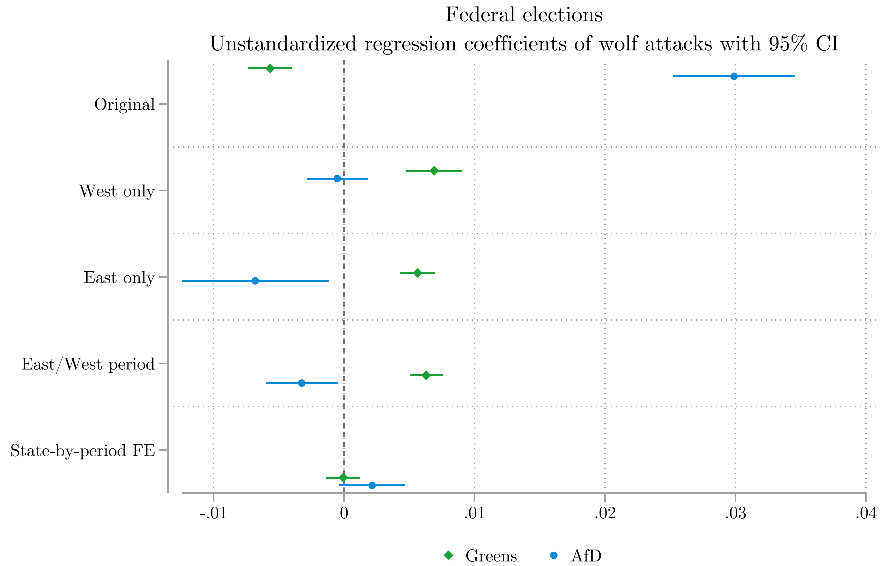

Fig. 5. Shown are point estimates and their 95% CIs from models regressing the AfD's (blue)/the Green party's (green) vote share on a dummy indicating that a municipality experienced a wolf attack. CIs are based on standard errors clustered at the municipality level. Data: Clemm von Hohenberg and Hager (2022); own calculation.

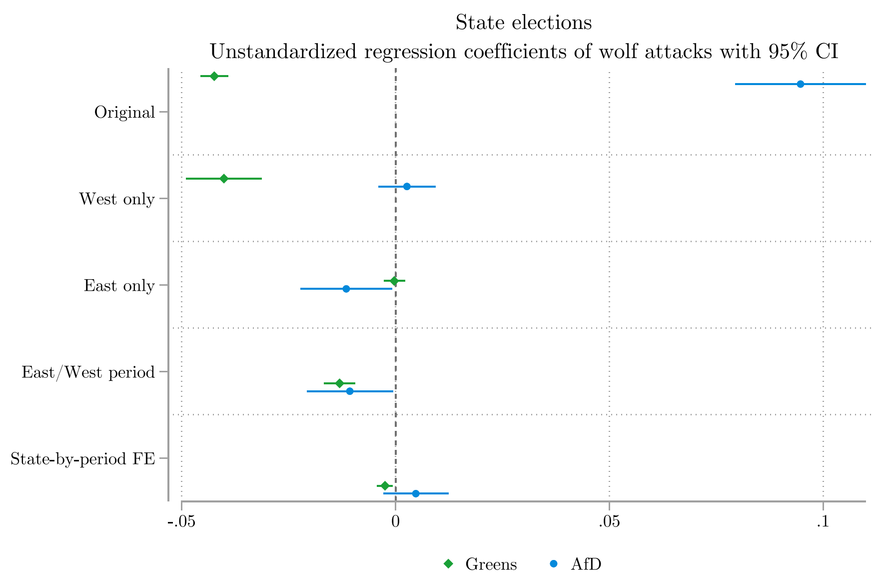

Fig. 6. Shown are point estimates and their 95% CIs from models regressing the AfD's (blue)/the Green party's (green) vote share on a dummy indicating that a municipality experienced a wolf attack. CIs are based on standard errors clustered at the municipality level. Data: Clemm von Hohenberg and Hager (2022); own calculation.

elections but negative for state elections. Only specifications with state-by-period fixed effects consistently yield coefficient estimates close to zero. In any case, it is evident that diverging time trends in party votes confounded the original results. While it is open to debate whether my proposed specifications are sufficient to remove all sources of time-varying confounding, the original study design was clearly deficient. It did not even allow to draw preliminary conclusions about the directions of the hypothesized causal effects.

3.2. An improved difference-in-differences design

The previous model specifications intentionally adhered closely to the original design in order to demonstrate the possible influence of divergent regional trends. However, the TWFE as implemented has at least three additional problems: First, wolf attacks happened at different points in time. Therefore, the treatment is "staggered" and there are different "treatment cohorts", i.e. some municipalities received the treatment in 2013, others in 2017 or 2021, etc. If the magnitude of the effect varies with treatment timing, then the estimate of the average across treatment cohorts might be hard-to-interpret due to the problem of "forbidden comparisons" (Roth et al., 2023, 2219). However, heterogeneous treatment effects can be modeled and averaged in a meaningful way (Roth et al., 2023, 2227-2228). Second, the static specifications in Eqs. (1)–(2) do not model how the treatment effect unfolds over time. A causal estimand that implicitly averages over all post-treatment periods might be difficult to interpret. Lastly, the models estimating the impact of wolf attacks on the AfD vote share contain some "always treated" units, that is, units that were already treated in 2013 and for which there are no pre-treatment observations. These units are neither a proper treatment group since they do not switch treatment state nor a proper control group because they have already been exposed to the treatment.

Focusing on the AfD vote share in federal elections, an improved DiD design should be based on three observations (elections) per unit:

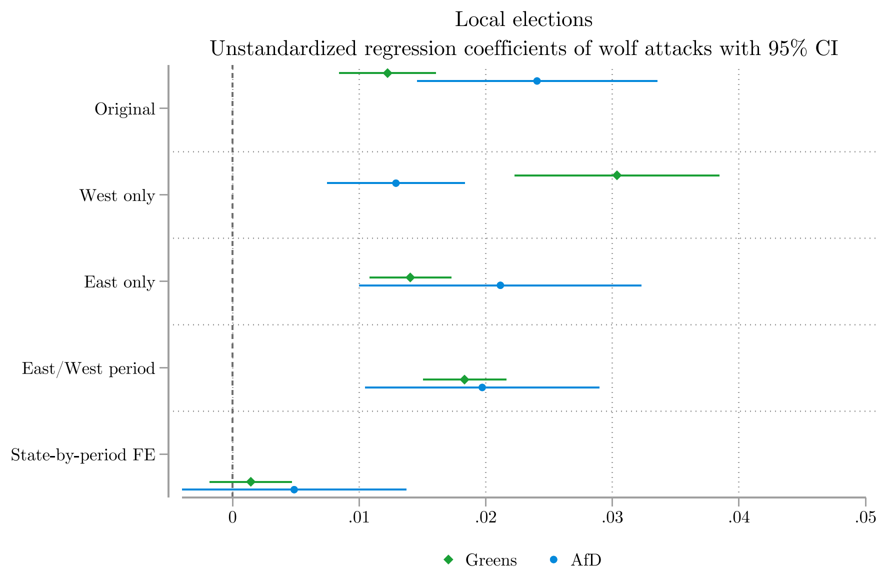

Fig. 7. Shown are point estimates and their 95% CIs from models regressing the AfD's (blue)/the Green party's (green) vote share on a dummy indicating that a municipality experienced a wolf attack. CIs are based on standard errors clustered at the municipality level. Data: Clemm von Hohenberg and Hager (2022); own calculation.

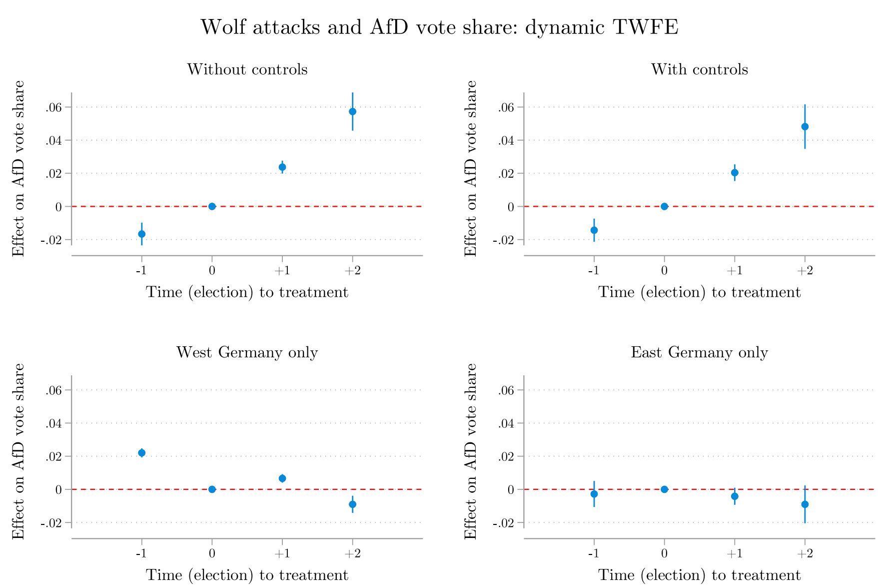

Fig. 8. Dynamic regression specification. Shown are point estimates and their 95% CIs based on standard errors clustered at the municipality level. Data: Clemm von Hohenberg and Hager (2022); own calculation.

2013, 2017, 2021. All municipalities that experienced wolf attacks before 2017 should be dropped from the sample in order to ensure that all remaining units are untreated in 2013. The control group should be comprised exclusively of "never treated" municipalities where no wolf attacks have taken place until after the election in 2021. Moreover, the treatment should differentiate between units treated in 2017 and units treated in 2021 to account for potential treatment cohort heterogeneity. Lastly, the short observation span (three periods) limits the possibility to implement an event study to capture treatment dynamics as well as possible pre-treatment trends. However, it is possible to model how the treatment effects unfolds over time for units treated in 2017 to see whether the effects persists in 2021. Likewise, for units treated in 2021 it is possible to estimate divergent pre-treatment trends.

The dynamic TWFE variant is based on the regression equation

$$\text{vote}_{it} = \alpha_i + \beta_t + \delta_{-1} D_{-1it} + \delta_1 D_{1it} + \delta_2 D_{2it} + \varepsilon_{it}. \tag{4}$$

It is similar to Eq. (1) except that the treatment indicator has been replaced by a set of dummy variables referring to relative treatment timing, i.e. D_{-1} equals 1 for municipalities two periods before the first wolf attacks takes place, D_1 indicates the first and D_2 the second post-treatment period. The period immediately preceding the treatment is the reference level. Therefore, the indicator D_0 is missing from the equation. Please keep in mind that $D_{-1} = 1$ can only be the case for units treated in 2021 and $D_2 = 1$ is possible only for units treated in 2017. This specification is valid if there is no effect heterogeneity with regard to treatment cohort. However, if voters of municipalities that first experienced a wolf attack prior to the 2021 election reacted differently compared to those of municipalities that experienced an attack in between the 2017 and the 2021 elections — e.g. due to changes in the general political climate — then the weighted mean across both groups as estimated by δ_1 can be difficult to interpret because the implicit weights of the cohort-specific effects might not be substantively meaningful. Hence, the recent literature recommends to model the cohort-specific effects and then average according to an explicit rule. I follow Sun and Abraham (2021) and weight according to the sample proportion of the treatment cohorts.[10] In total, I present results for four specifications: First, I estimate the model without and with adjustment for the four time-varying covariates used by Clemm von Hohenberg and Hager (2022).[11] Second, I estimate the model separately for East and West German subsamples.

The results in Fig. 8 are in line with the findings in Section 3.1. The pooled analyses in the upper column are consistent with a divergence between municipalities susceptible to wolf attacks and other regions. Adjusting for the number of refugees, unemployment and land use results in no improvement. In fact, the trend remains surprisingly linear. Splitting the sample by region in general seems to confirm previous results as well: no indication of a pre-treatment trend and a small, but insignificant negative coefficient among East German municipalities; some indication of pre-treatment trends and no clear associations between wolf attacks and AfD vote share post-treatment in West Germany.

4. Just when I thought theory was out, they pull it back in

Clemm von Hohenberg and Hager (2022) use a variant of the popular TWFE estimator which — under certain assumptions — implements a generalized difference-in-differences design (Roth et al., 2023). TWFE is often the default empirical strategy and seems preferable to more structural, theory-guided approaches due to its promise to eliminate unit- and period-specific errors with minimal parametric assumptions. However, it has been pointed out in methodological literature that TWFE relies on assumptions that are far from being "harmless" (Plümper and Troeger, 2019; Imai and Kim, 2021). More specifically, as soon as we leave the rather narrow set of true natural experiments and adjustment for time-varying confounders or the modeling of time-varying unobserved heterogeneity seems warranted, we find ourselves once again in a situation where we are confronted with an overwhelming number of inevitable decisions about functional forms and other fraught issues.

For example, Clemm von Hohenberg and Hager (2022) adjust for the number of refugees per municipality and year. This might or might not be the correct functional specification since there are equally plausible alternatives. In order to understand support for anti-immigration parties, it could be important to take the "initial conditions" of recent developments into account. For instance, a previously ethnically homogeneous community, compared an already diverse community, might react differently to the same absolute increase in the number of refugees. Therefore, it could be more appropriate to use an indicator of relative rather than absolute change of the refugee population (or immigrants in general).[12]

However, the empirical implementation is not as straightforward as it might seem. In the data at hand, modeling relative changes in the refugee population is complicated by the presence of many initial zeros. Hence, it is not possible to just divide the number of refugees in a municipality in a given year by, for example, the variable's value in 2009 — the last federal election without AfD participation — because this would result in losing most observation due to missing values. What is more, the interaction between an observable covariate, like number of refugees, and unit-specific unobserved factors, e.g. conditions of the local culture and community prior to recent demographic changes, could be more complicated than can be captured by a simple transformation.

Fixed-effect individual slopes (FEIS) offer a possible solution to both problems. FEIS are most often used to model unit-specific time trends but this is just a special case of a general class of data-generating processes (Rüttenauer and Ludwig, 2023; Brüderl and Ludwig, 2014). Similar to random slopes, it is possible to estimate FEIS for any covariate (conditional on data requirements). For example, instead of estimating

$$\text{vote}_{it} = \alpha_i + \beta_t + \gamma \text{refugees}_{it} + \delta \text{wolf}_{it} + \varepsilon_{it} \tag{5}$$

where refugees_{it} is the number of refugees in municipality i at election t, we could estimate an FEIS model with varying slopes

$$\text{vote}_{it} = \alpha_{0i} + \alpha_{1i} \text{refugees}_{it} + \beta_t + \delta \text{wolf}_{it} + \varepsilon_{it}. \tag{6}$$

The strength of the response to a change in the number of refugees (α_{1i}) now varies across municipalities. Of course, there are parametric choices implicit in FEIS as well. For example, in Eq. (6), FEIS are assumed to be linear and the relationship to be symmetric, i.e. an increase in the number of refugees has the same "effect" as an decrease of the same size (for general remarks on this often implicit assumption: Haffert and Ergen, 2019; Allison, 2019).

The preceding remarks sketch some of the degrees of freedom available but it is beyond the limits of this research note to explore whether they bear any actual relevance for explaining divergent trends in vote shares among municipalities in different German regions. The point is that when simple design-based inference fails due to a lack of quasi-random assignment, the task of adjusting for heterogeneous trends is unlikely to be accomplished without theorizing about the underlying mechanisms.

5. Conclusion

There are several lessons to be drawn from this note. Substantively, it seems doubtful whether the local reactions to wolf attacks are powerful enough to influence the election decisions of many voters. Attempting to correct for regional trends in voting reverses the sign — far-right voting decreases, Green Party votes increase — or eliminates the relationship altogether, but all in all the current research designs is not suitable to answer the research question with certainty. The large effect sizes reported in the original study alone should make researchers cautious.

Methodologically, the issue of region-specific time trends illustrates the pitfalls of spatial panel data (Auspurg et al., 2019). The authors transparently plotted the spatial distribution of the treatment in their original article but did not draw any conclusions about the possible problems resulting from the non-random distribution. At a bare minimum, simple descriptive evaluations of the main identification assumptions should have been performed (Section 2). I have refrained from using the toolkit of spatial statistics to simplify things but it would have been possible to assess the spatial correlations of residuals more systematically. For example, divergent pre-treatment trends could result from spill-over effects in voter sentiment from neighboring municipalities because early-treated and late-treated units cluster geographically. Furthermore, some of the robustness checks proposed by Kelly (2020) for historical persistence research, such as adjusting for

[10] Estimation is based on the procedure implemented by Sun (2021).

[11] The controls include the number of refugees, the proportion unemployed, the proportion of area used for agriculture, and the proportion classified as heath. The last covariate is a bit of an odd choice. The authors state they have "controlled for common determinants of far-right voting" (Clemm von Hohenberg and Hager, 2022, 3). However, I am unaware of any research showing that changes in scrubland habitats influence far-right voting, though it could be linked to wolf migration. (The variable is labeled "prop_heathen" in the dataset and, indeed, adjusting for the proportion of atheists instead would make more sense theoretically.)

[12] The importance of relative rather than absolute changes could be derived from a kind of "(reverse) contact hypothesis". For instance, individuals who grew up or lived in diverse major cities during recent decades may give less weight to the potential benefits of living in an ethnically homogeneous community than those who, until recently, lived in very homogeneous communities. Related ideas have been discussed as "compositional amenities" of ethnic homogeneity, with some research finding that related concerns may explain most of the rural–urban gap in attitudes toward immigration (Card et al., 2012, 96).

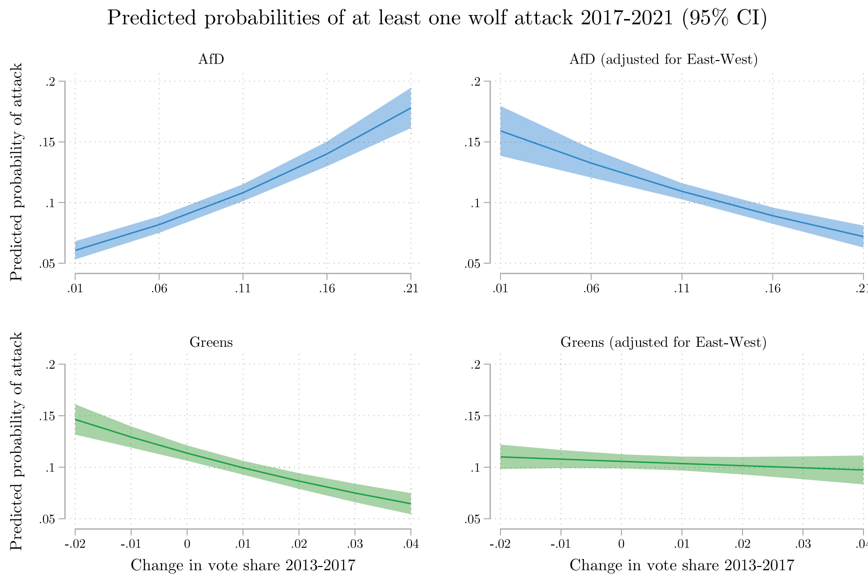

Fig. A.9. Probability of a wolf attack during the period 2017–2021 as predicted by changes in the vote shares between 2013–2017. On the left, predictions are based on unadjusted bivariate probit regressions. On the right, predictions are based on estimates adjusted for a regional dummy covariate. The range of the predictor is restricted to the interval between its 5th and 95th percentile. Data: Clemm von Hohenberg and Hager (2022); own calculation.

longitude and latitude, could be valuable for aggregate spatial analyses in general. This holds true even when the data and events under investigation are recent, and the observational time-span is relatively short. Studies like Clemm von Hohenberg and Hager (2022) adopt a similar approach to persistent research, focusing on the lasting impact of one-time events that vary in their geographic distribution. Lastly, as discussed in Section 4, there are few "ready-made designs". The admission that treatment distribution is non-random and, thus, time-varying controls may be necessary, implies theory-guided reflections on the appropriate modeling choices. The adoption of methods as propagated in the context of the "Credibility Revolution" should be accompanied by a stronger integration with theory development (Ashworth et al., 2021). Still, theories are frequently insufficient to give exact guidance. In this case, systematic robustness analyses or a transparently descriptive overall framing might be more appropriate. Descriptive research is important and it is a pity that the general obsession with causal estimates disincentivizes researchers from attempting to publish careful and detailed description.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I thank Bernhard Clemm von Hohenberg and Anselm Hager for their openness in providing information and for a constructive discussion regarding my concerns. I thank Nicole Schwitter, Dave Balzer, Daniel Meyer and three anonymous reviewers for helpful feedback on earlier drafts of this research note.

Appendix. Far-right voting predicts wolf attacks

The most obvious (observable but neglected) characteristic driving the association between prior vote share and wolf attacks might be the location of the municipality. As mentioned in the introduction, wolf attacks are much more likely to happen in East Germany. A simple analysis illustrates the possible confoundedness of the treatment assignment. In a first step, I estimate a bivariate probit regression (subscript i dropped for simplicity)

$$P(\text{wolf}_{2021} = 1|\text{vote}_{2017-2013}) = \Phi(\alpha + \beta\text{vote}_{2017-2013}) \quad \text{(A.1)}$$

where $\Phi()$ is the cumulative normal distribution function, wolf_{2021} equals 1 if there was a wolf attack in a municipality in the period before the last federal election 2017–2021, and the sole predictor is $\text{vote}_{2017-2013}$, i.e. the change in the vote share between the two previous elections (2013 to 2017). In a second step, I adjust for East-West differences by additionally including a dummy variable for the region.

The left column in Fig. A.9 shows predicted probabilities based on two separate bivariate models for AfD and Green Party changes. A growing AfD vote share between 2013 and 2017 predicts a higher probability of a wolf attack in the period from 2017 to 2021. *Vice versa*, the greater the change in the Green Party vote share, the lower the predicted probability of a wolf attack. Adjusting for East-West difference by including a dummy variable for East Germany and setting its value to its mean eliminates (Greens) or inverts (AfD) the relationship between pre-treatment voting trends and wolf attacks (Fig. A.9, right column). Thus, trends in far-right/Green voting for the period 2013–2017 predict wolf attacks for the period 2017–2021 partly because wolf attacks are more common in East Germany.

Data availability

All data, syntax/code, and a brief README with instructions are openly available on the Open Science Framework (OSF): https://osf.io/6uvnf/.

References

Allison, P., 2019. Asymmetric fixed-effects models for panel data. Socius 5, 2378023119826441. http://dx.doi.org/10.1177/2378023119826441.

Andersen, L.W., Harms, V., Caniglia, R., Czarnomska, S.D., Fabbri, E., Jędrzejewska, B., Kluth, G., Madsen, A.B., Nowak, C., Pertoldi, C., et al., 2015. Long-distance dispersal of a wolf, canis lupus, in northwestern europe. Mammal Res. 60, 163–168. http://dx.doi.org/10.1007/s13364-015-0220-6.

Arzheimer, K., 2023. The electoral breakthrough of the AfD and the east-west divide in German politics. In: Weisskircher, M. (Ed.), Contemporary Germany and the Fourth Wave of Far-Right Politics. Routledge, London, pp. 140–158. http://dx.doi.org/10.4324/9781003120049-10.

Ashworth, S., Berry, C.R., de Mesquita, E., 2021. Theory and Credibility: Integrating Theoretical and Empirical Social Science. Princeton University Press, Princeton, http://dx.doi.org/10.23943/princeton/9780691213828.001.0001.

Auspurg, K., Brüderl, J., Wöhler, T., 2019. Does immigration reduce the support for welfare spending? A cautionary tale on spatial panel data analysis. Am. Sociol. Rev. 84 (4), 754–763. http://dx.doi.org/10.1177/0003122419856347.

Brodeur, A., Esterling, K., Ankel-Peters, J., Bueno, N.S., Desposato, S., Dreber, A., Genovese, F., Green, D.P., Hepplewhite, M., de la Guardia, F.H., Johannesson, M., Kotsadam, A., Miguel, E., Velez, Y.R., Young, L., 2024. Promoting reproducibility and replicability in political science. Res. Politics 11 (1), 20531680241233439. http://dx.doi.org/10.1177/20531680241233439.

Brüderl, J., Ludwig, V., 2014. Fixed-effects panel regression. In: Best, H., Wolf, C. (Eds.), The SAGE Handbook of Regression Analysis and Causal Inference. SAGE, London, pp. 327–357. http://dx.doi.org/10.4135/9781446288146.n15.

Card, D., Dustmann, C., Preston, I., 2012. Immigration, wages, and compositional amenities. J. Eur. Econ. Assoc. 10 (1), 78–119. http://dx.doi.org/10.1111/j.1542-4774.2011.01051.x.

Clemm von Hohenberg, B., 2022. Replication data for: Wolf attacks predict far-right voting. http://dx.doi.org/10.7910/DVN/OOHLYX.

Clemm von Hohenberg, B., Hager, A., 2022. Wolf attacks predict far-right voting. Proc. Natl. Acad. Sci. 119 (30), e2202224119. http://dx.doi.org/10.1073/pnas.2202224119.

Clemm von Hohenberg, B., Hager, A., 2024a. Correction for Clemm von Hohenberg et al., Wolf attacks predict far-right voting. Proc. Natl. Acad. Sci. 121 (33), e2412545121. http://dx.doi.org/10.1073/pnas.2412545121.

Clemm von Hohenberg, B., Hager, A., 2024b. Wolf attacks predict far-right voting. Replication guide & supplementary materials (corrected version). http://dx.doi.org/10.7910/DVN/OOHLYX.

Cochrane, J.H., 2018. A brief parable of over-differencing. Unpublished Manuscript. Chicago: University of Chicago, Booth School of Business. https://www.johnhcochrane.com/s/overdifferencing.pdf.

Correia, S., 2016. A feasible estimator for linear models with multi-way fixed effects. Preprint at http://scorreia.com/research/hdfe.pdf.

Correia, S., 2023. REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. GitHub. https://github.com/sergiocorreia/reghdfe.

Haffert, L., Ergen, T., 2019. The symmetric fallacy: the dangers of symmetric reasoning in the social sciences. In: CES Open Forum Series. Center for European Studies, Harvard University, https://issuu.com/ces.harvard/docs/working_paper_pdf_-_the_symmetric_f.

Imai, K., Kim, I.S., 2021. On the use of two-way fixed effects regression models for causal inference with panel data. Political Anal. 29 (3), 405–415. http://dx.doi.org/10.1017/pan.2020.33.

Imai, K., Kim, I.S., Wang, E.H., 2023. Matching methods for causal inference with time-series cross-sectional data. Am. J. Political Sci. 67 (3), 587–605. http://dx.doi.org/10.1111/ajps.12685.

Kelly, M., 2020. Understanding persistence. CEPR Discussion Paper No. 15246, CEPR Press, Paris & London, https://cepr.org/publications/dp15246.

Plümper, T., Troeger, V.E., 2019. Not so harmless after all: The fixed-effects model. Political Anal. 27 (1), 21–45. http://dx.doi.org/10.1017/pan.2018.17.

Roth, J., Sant'Anna, P.H., Bilinski, A., Poe, J., 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. J. Econometrics 235 (2), 2218–2244. http://dx.doi.org/10.1016/j.jeconom.2023.03.008.

Rüttenauer, T., Ludwig, V., 2023. Fixed effects individual slopes: Accounting and testing for heterogeneous effects in panel data or other multilevel models. Sociol. Methods Res. 52 (1), 43–84. http://dx.doi.org/10.1177/0049124120926211.

Schmidt-Catran, A.W., Spies, D.C., 2016. Immigration and welfare support in Germany. Am. Sociol. Rev. 81 (2), 242–261. http://dx.doi.org/10.1177/0003122416633140.

Schmidt-Catran, A.W., Spies, D.C., 2019. Immigration and welfare support in Germany: Methodological reevaluations and substantive conclusions. Am. Sociol. Rev. 84 (4), 764–768. http://dx.doi.org/10.1177/0003122419858729.

Sun, L., 2021. Eventstudyinteract: Interaction weighted estimator for event study. GitHub. https://github.com/lsun20/eventstudyinteract.

Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J. Econometrics 225 (2), 175–199. http://dx.doi.org/10.1016/j.jeconom.2020.09.006.

Weisskircher, M., 2020. The strength of far-right AfD in Eastern Germany: The east-west divide and the multiple causes behind 'Populism'. Political Q. 91 (3), 614–622. http://dx.doi.org/10.1111/1467-923X.12859.

Zscheischler, J., Friedrich, J., 2022. The wolf (canis lupus) as a symbol of an urban–rural divide? Results from a media discourse analysis on the human–wolf conflict in Germany. Environ. Manag. 70 (6), 1051–1065. http://dx.doi.org/10.1007/s00267-022-01719-3.