Public Service Failures and Voter Participation: Evidence from Court-Validated Electoral Disruptions*

Marius Kröper†

July 7, 2025

Preliminary Draft — Please do not cite or distribute

Abstract

This paper studies how administrative failures on election day affect long-run political participation. I exploit court-validated disruptions during Berlin's 2021 elections—such as ballot shortages and excessive queues—as a natural experiment. Using an event-study design across nine elections, I show that turnout in affected precincts fell by 1.2 percent and remained depressed for at least 4 years. Effects are largest among young voters and residents with migration backgrounds. Survey evidence suggests both trust erosion and disrupted civic habits. The findings underscore that competent electoral administration is essential for democratic inclusion and should be treated as core public infrastructure.

Keywords: Postal Voting, Voter Turnout, Elections, Administration, Disruptions, Berlin, Voting Costs

JEL-Codes: D72, H11, H70, R50

*I gratefully acknowledge the invaluable feedback from Alexander Kemnitz, Enrico Cantoni, Valentin Lindlacher, Giacomo Opocher, and the participants of the TUD Brown Bag Seminar and the RGS 2025 conference.

†TUD Dresden University of Technology. marius.kroeper@tu-dresden.de. Helmholtzstr. 10, 01069 Dresden (Germany).

1 Introduction

Elections are not only expressions of democratic choice — they are also a form of public service delivery. Administered by the state, they require reliable infrastructure, competent logistics, and equitable access. Like schools or health systems, the quality with which elections are delivered shapes how citizens interact with public institutions. When these processes function smoothly, they help build civic habits and sustain trust in the state (Fujiwara et al., 2016). But when they break down—through ballot shortages, excessive queues, or administrative confusion—they risk undermining that trust and discouraging future participation (Fujiwara et al., 2016; Pettigrew, 2021).[1]

Much of the existing literature on voting behavior emphasizes how formal rules and logistical frictions can restrict access to the ballot box. Voter ID laws, polling place relocations, reduced time to vote, or long travel distances are often policy-driven or institutionally sanctioned and have been shown to depress turnout, particularly among marginalized groups (Cantoni, 2020; Cantoni and Pons, 2021; Alipour and Lindlacher, 2025; Kaplan and Yuan, 2020; Potrafke and Roesel, 2020). This paper shifts attention from the design of electoral institutions to their delivery. It studies the consequences of unintentional administrative failures in electoral service delivery—failures that are operational rather than strategic, but nonetheless consequential. While such disruptions are infrequent in high-capacity democracies, they can reduce political participation and exacerbate existing inequalities. The analysis demonstrates that administrative breakdowns, such as ballot shortages or excessive queuing, can lead to persistent declines in turnout, particularly among groups that already face structural barriers to participation.

The 2021 Berlin elections provide a rare opportunity to study the consequences of large-scale administrative failure in a high-capacity democracy. Held on the same day as the *Berlin Marathon* and under the constraints of *COVID-19 pandemic protocols*, the election combined *four contests*—federal, state, district, and referendum—placing exceptional demands on local electoral infrastructure. Despite Germany's reputation for bureaucratic competence, hundreds of precincts experienced substantial disruptions: ballot shortages, incorrect ballot allocations, polling station closures, and waiting times exceeding three hours. These failures triggered multiple court proceedings. The *Berlin Constitutional Court* annulled the state election results statewide, while the *Federal Constitutional Court* mandated partial reruns of the federal election in 431 precincts. These rulings provide a unique empirical setting: treatment is assigned exogenously by judicial decision, based on documented administrative irregularities, rather than inferred from media or survey data, allowing causal identification of electoral misadministration on political behavior.

Voting is as a low-stakes, high-cost activity, where small frictions can shape long-term behavior through habit formation. In standard rational-choice models, individuals turn out when the expected benefits exceed these costs (Downs, 1957; Riker and Ordeshook, 1968). Administrative disruptions—such as long queues, ballot shortages, or procedural confusion—can elevate per-

[1] Alsan and Wanamaker (2018) show, for example, that reduced trust in medical providers led to long-run declines in health care utilization.

ceived costs in the short run and impede the formation of civic habits over time. This mechanism generates persistence: voters exposed to elevated costs in one election may remain disengaged in subsequent cycles, even after the original disruption has been resolved. Fujiwara et al. (2016) show that adverse conditions in voters' first eligible election reduce turnout in later contests, consistent with habit-formation dynamics. In the Berlin case, affected voters may have revised their expectations of electoral reliability, reducing their long-term propensity to participate.

Beyond habit disruption, administrative failures may also affect turnout by altering beliefs about state competence. When voters encounter visible mismanagement—such as ballot shortages, procedural confusion, or long queues—they may revise upward their expectations of future participation costs or downward their confidence in electoral integrity. These informational shocks are likely to matter most for marginal voters, whose beliefs about procedural reliability are less stable. In high-capacity democracies, where expectations of bureaucratic competence are strong, such failures may be particularly salient. A growing literature shows that trust in public institutions is sensitive to delivery failures: Alsan and Wanamaker (2018) document persistent health care avoidance among Black men following the disclosure of the Tuskegee Study; Lowes and Montero (2021) find that coercive colonial medical campaigns reduced trust and utilization of health services; and Gottlieb (2016) shows that unmet expectations about local service delivery undermine citizen confidence in government. In the Berlin case, even unintended administrative failures may have served as a signal of state fragility, reinforcing disengagement beyond the immediate disruption.

The empirical analysis exploits quasi-experimental variation induced by judicially mandated reruns of the 2021 Berlin elections. Following documented irregularities—including ballot shortages, incorrect ballot deliveries, and excessive wait times—the *Berlin Constitutional Court* annulled the state-level election citywide, and the *Federal Constitutional Court* ordered a partial rerun of the federal election in 431 precincts. Treatment assignment is therefore exogenous, based on legal rulings rather than self-reports or administrative discretion.

The estimation sample covers nine elections between 2014 and 2025. Turnout is measured at the level of harmonized postal precincts, constructed using spatial raster weighting to address changes in precinct boundaries over time. The main specification is a stacked event-study design, which estimates dynamic and heterogeneous treatment effects across electoral cycles. Outcomes include total turnout, vote mode, and party-level vote shares.

To complement the administrative analysis, I incorporate data from the German Longitudinal Election Study (GLES). These survey data provide individual-level information on political trust, perceptions of procedural fairness, and abstention timing. While not used for identification, the survey outcomes offer corroborating evidence on the mechanisms linking administrative failure to political disengagement.

Electoral service failures in Berlin's 2021 elections led to persistent and demographically uneven declines in voter turnout. Treated precincts experience an average reduction in log turnout of approximately 1.2 percent, with no evidence of recovery over three years or across subsequent electoral cycles. The decline is driven by sustained reductions in in-person voting,

only partially compensated by a gradual increase in postal turnout. Even in the long run, total turnout in affected precincts remains significantly below pre-treatment levels.

The effects are heterogeneously concentrated. Turnout declines are most pronounced in precincts with a higher share of young voters (ages 18–35), residents with a migration background, and welfare recipients. These patterns are consistent with models in which marginal voters respond more elastically to increases in participation costs or signals of administrative failure. Estimated treatment effects are also larger in precincts that experienced longer wait times during the 2021 disruption.

Survey data from the German Longitudinal Election Study support the interpretation of both short-term disruption and long-run disengagement. In the immediate aftermath of the election, respondents in Berlin report significantly lower trust in electoral fairness. While institutional trust partially recovers, the reported incidence of last-minute abstention is substantially elevated, indicating disrupted decision-making at the point of participation.

Electoral consequences at the party level are limited but measurable. Vote shares for most parties remain stable, but treated precincts exhibit a small but statistically significant decline in support for the *CDU* (conservative) and a corresponding increase in *FDP* (liberal) vote share. These shifts suggest that administrative failures may have distributional effects on party competition, particularly among swing or protest voters.

This paper contributes to three strands of literature.

First, it provides new evidence that administrative failures in electoral service delivery can generate persistent reductions in political participation. Prior work has documented short-run effects of legal or logistical voting barriers—including polling station relocations (Alipour and Lindlacher, 2025), voter ID laws (Cantoni and Pons, 2021), participation rules (Gerber et al., 2003; Potrafke and Roesel, 2020; Bechtel et al., 2018; Hoffman et al., 2017; Gaebler et al., 2020), and distance to polling locations (Cantoni, 2020)—but most of these effects attenuate over time. A notable exception is Pettigrew (2021), who documents persistent turnout declines following exposure to restrictive ID laws. This paper differs in using judicial rulings to define treatment, eliminating perception bias and ensuring exogeneity. It thus adds to recent efforts to credibly identify the consequences of electoral disruptions (Cantoni et al., 2025).[2]

Second, the paper integrates habit formation and informational updating into a unified behavioral framework for civic participation. Administrative failures may affect voting both by disrupting civic routines and by degrading beliefs about state competence. This builds on work showing long-run behavioral scarring from early voting experiences (Fujiwara et al., 2016) and trust erosion from exposure to public service failure (Alsan and Wanamaker, 2018; Lowes and Montero, 2021; Gottlieb, 2016; Rosas, 2010).

Third, by framing electoral disruptions as public service failures, the paper connects political economy models of turnout with broader research on bureaucratic performance and state capacity. Failures in election delivery resemble dysfunctions in education, health, or policing—domains where the quality of implementation shapes trust and engagement (Akhtari et al.,

[2] For an overview, see Cantoni et al. (2025).

2022; Toral, 2023). The findings thus build a conceptual bridge between the study of democratic participation and the literature on bureaucratic accountability and service delivery under stress.

The results highlight that electoral integrity depends not only on legal safeguards or fraud prevention, but also on the reliability of administrative delivery. Disruptions in routine electoral services can suppress participation—especially among already underrepresented groups—even in high-capacity democracies. These findings suggest that resilient election logistics are essential for democratic inclusion, and that administrative competence should be treated as a core component of electoral infrastructure.

The rest of the paper proceeds as follows. The setting and the data are described in Section 2, followed by theoretical considerations and the empirical strategy in Section 3. The main results are presented in Section 4, while results on party outcomes are shown in Section 6. Section 7 concludes.

2 Institutional Background and Data

Modern elections are not only exercises in democratic choice but also complex logistical undertakings. Their effectiveness as public services hinges on reliable administration. This section summarizes the institutional and empirical context for the Berlin 2021 election failures, the subsequent legal rulings, and the data infrastructure used to study their effects.

Berlin's electoral system follows Germany's mixed-member proportional model, combining district-level candidate votes (*Erststimme*) and party-list votes (*Zweitstimme*). On 26 September 2021, four elections were held concurrently—federal, state, district, and a city-wide referendum—amid ongoing Covid-19 precautions and the Berlin Marathon. These factors placed extreme stress on local electoral logistics.

The result was a cascade of irregularities: ballot shortages, distribution errors, delayed openings, premature closings, and queues exceeding two hours. Traffic disruptions from the marathon further impaired access to polling places. These breakdowns were not merely anecdotal. Judicial reviews by the Berlin and Federal Constitutional Courts found that the elections in hundreds of precincts violated basic requirements for orderly and equal electoral access. The Berlin Constitutional Court ordered a full rerun of the state election in February 2023. The Federal Constitutional Court later mandated a partial rerun of the federal election in 431 precincts, held in February 2024.

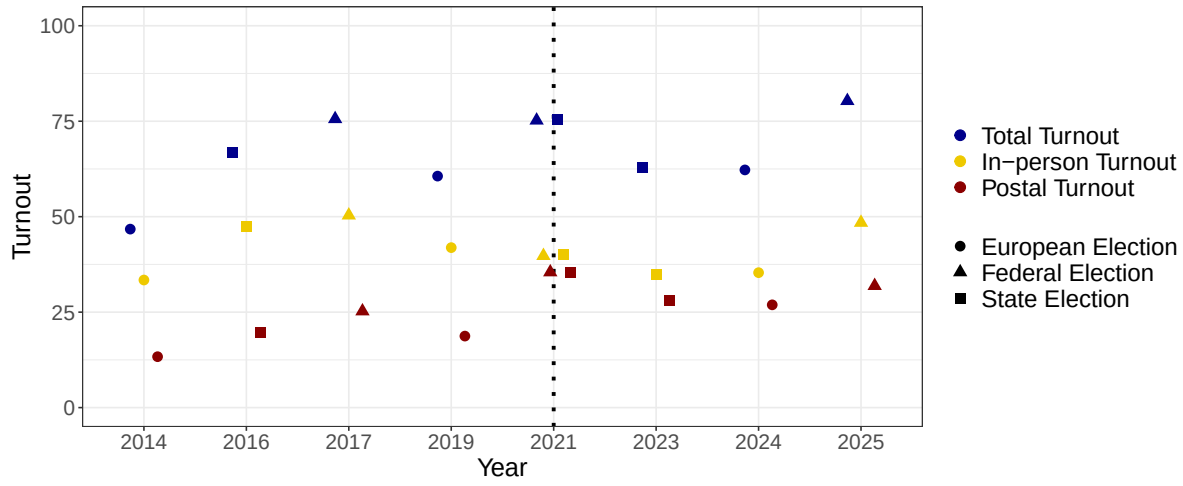
These rulings offer a rare source of court-validated, precinct-level treatment assignment. Unlike perception-based approaches relying on survey recall or media reports, this setting defines treatment exogenously: treated precincts are those the courts deemed to have experienced legally consequential misadministration.

The irregularities constitute a multidimensional disruption to voting: ballot unavailability and long queues raise participation costs; delays and confusion reduce procedural trust; and road closures impose physical access frictions. These mechanisms plausibly affect both whether and how people vote. Prior research shows that even modest logistical barriers—such as minor changes in polling station distance (Cantoni, 2020; Alipour and Lindlacher, 2025)—can suppress

turnout. The Berlin case involves larger, unplanned shocks, offering a test of behavioral and informational mechanisms in a high-capacity setting.

Germany's electoral infrastructure facilitates postal voting, which has expanded steadily since 2008. Voters can request ballots without cause and return them by mail or deposit. As shown in Figure 1, postal turnout rose sharply after 2014 and now comprises a substantial share of participation. This offers a substitute margin for those discouraged from in-person voting. However, substitution is incomplete: digital access, registration reliability, and information constraints can limit adoption.

Figure 1: Turnout by Election Type in Berlin since 2014



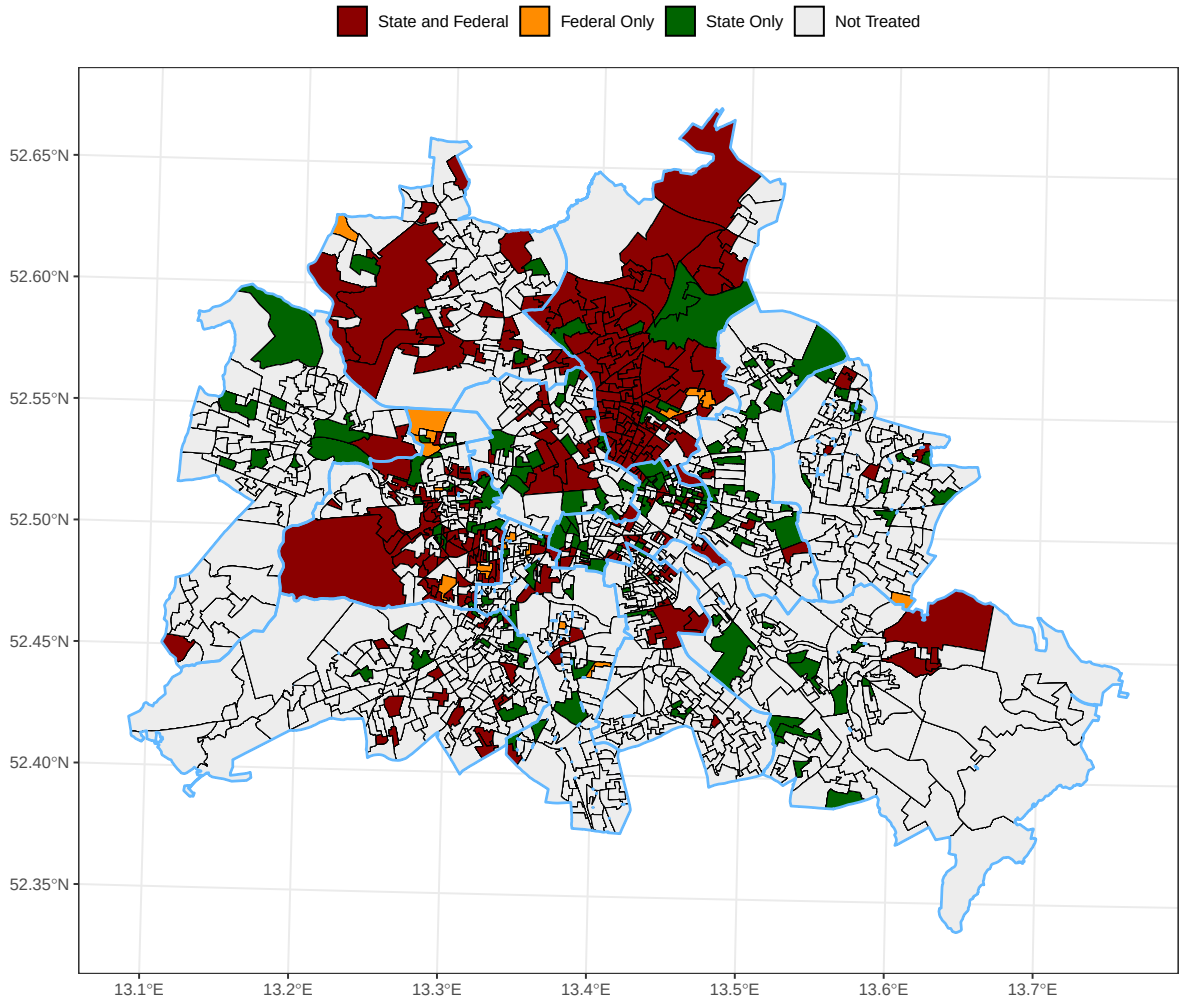
Notes: The figure presents total (yellow), in-person (blue), and postal turnout (green) for the European (circle), Federal (triangle), and State (square) elections in the sample. The dotted vertical line marks the 2021 election, which triggered the constitutional review.

The empirical design links these disruptions to long-run turnout effects using precinct-level panel data. The unit of analysis is the postal precinct (Briefwahlbezirk), the smallest unit reporting turnout by mode and party. Electoral data are harmonized to the 2021 precinct geography using population-weighted crosswalks based on 100-meter census grids. I exclude precincts affected only by the state court ruling and those with ballot misallocation issues. The final dataset spans eight elections over six years, with 431 precincts identified as treated by the Federal Court.

Figure 2 shows the spatial distribution of affected precincts. Treatment clusters in central districts but appears across the city.

The next section outlines the event-study design used to estimate both average and dynamic treatment effects across elections.

Figure 2: Berlin Precincts by Treatment



Notes: This figure shows postal precincts affected by election irregularities during the 2021 federal (orange), state (green), or both (red) elections. Districts are outlined in blue, 2021 postal precinct boundaries in black. Precincts in Tempelhof-Schöneberg are excluded due to inconsistent ballot allocation.

3 Theoretical Considerations and Empirical Framework

3.1 Behavioral Mechanisms: Habit Formation and Belief Updating

To interpret the persistent effects of electoral service disruptions, this section outlines a behavioral model integrating two mechanisms: *habit formation* and *informational updating*. The framework extends canonical models of voting under participation costs (Downs, 1957; Riker and Ordeshook, 1968) by allowing both past voting behavior and administrative signals to shape future turnout.

Consider a representative voter who decides whether to vote in two consecutive elections, $t = 0, 1$. In each period, the utility from voting is:

$$U_t = p_t B + d_t - c_t, \tag{1}$$

where B is the expressive or instrumental benefit from electoral participation, p_t is the subjective probability of pivotality, d_t captures intrinsic motivation or civic habit, and c_t is the perceived cost of voting. The voter chooses $v_t = 1$ if $U_t > 0$, and abstains otherwise.

The *habit formation* channel assumes civic engagement is path-dependent. Following Fujiwara et al. (2016), let:

$$d_1 = \bar{d} + \rho v_0, \tag{2}$$

where $\bar{d}$ is baseline intrinsic motivation and $\rho > 0$ governs habit formation. If a service disruption at $t = 0$ increases costs by δ, such that $c_0 = \bar{c} + \delta$, it reduces v_0, which in turn lowers d_1 and depresses future participation.

The *informational updating* channel captures the impact of observed service quality on expectations. Voters observe a noisy signal $s = \theta + \varepsilon$, where θ is the latent administrative quality and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. This signal informs expectations about future voting costs:

$$c_1 = \bar{c} + \phi(s), \quad \text{with } \phi'(s) > 0. \tag{3}$$

If a salient disruption occurs—e.g., ballot shortages or long queues—then s is low, raising $\phi(s)$ and thus increasing c_1. This reduces the probability of turnout in $t = 1$, particularly for voters with weak priors or low institutional trust.

Combining both channels, the probability of voting in $t = 1$ is:

$$P(v_1 = 1) = \Pr(p_1 B + \bar{d} + \rho v_0 - (\bar{c} + \phi(s)) > 0). \tag{4}$$

This formulation highlights two complementary behavioral forces: prior participation increases civic habit, while perceived mismanagement increases expected costs. Both mechanisms reduce future participation.

The model implies persistent turnout decline, stronger effects among first-time voters or low-trust groups, and incomplete recovery even after administrative quality improves. This is because habits, once broken, may take multiple election cycles to reestablish, particularly for marginal voters. The model also explains why substitution to mail voting may be partial: if disruptions undermine trust in the electoral process as a whole, voters may generalize these beliefs across voting modes, reducing participation even where logistical costs are lower.

While the *habit formation* and *informational updating* mechanisms are analytically distinct, they are jointly triggered by the treatment in this setting and cannot be separately identified in the data. The empirical analysis provides suggestive evidence for both by exploiting heterogeneity in prior voting behavior, institutional trust, and the timing of abstention decisions.

3.2 Estimation Equation

To assess these predictions empirically, I estimate a dynamic difference-in-differences model using a stacked event study design. The setting is well-suited to this approach: the treatment is assigned at the postal precinct level, all treated units were exposed simultaneously in 2021,

and data are available across multiple election cycles for the same spatial units. The following equation forms the basis of the estimation strategy:

$$Y_{idt} = \sum_{\tau \neq -1} \beta^{\tau} (\mathbb{1}_{\tau=t} \times \text{Treatment}_i) + \alpha_i + \alpha_{dt} + \varepsilon_{idt}, \tag{5}$$

where Y_{idt} denotes the voting outcome of interest——such as the natural log of in-person turnout, postal turnout, or total turnout——in postal precinct i, district d, and election year t. The variable Treatment_i is an indicator equal to one for precincts that were formally declared affected by administrative failures during the 2021 federal election by the Federal Constitutional Court. The term $\mathbb{1}_{\tau=t}$ denotes an indicator for each election year relative to the disruption year, where $\tau - 1$ corresponds to the pre-treatment baseline (the 2019 European Parliament election), and $\tau \in \{-4, -3, -2, 0/F, 0/S, 1, 2, 3\}$ includes the elections before and after the disruption.[3] The fixed effect α_i accounts for time-invariant heterogeneity across precincts, while α_{dt} captures any election-year shocks or trends that vary at the district level. Standard errors are clustered at the precinct level to address serial correlation across elections within units.

The coefficients β^{τ} capture the dynamic treatment effects of exposure to administrative failure, tracing the evolution of voting behavior before and after the disruption. A negative and significant coefficient in the election year 2021 $(t + 0)$ would suggest that disruptions causally suppressed turnout at the time of the election. If subsequent coefficients remain non-zero, this would provide evidence that the behavioral effects of the failure persisted over time, either through learned substitution, reduced institutional trust, or disrupted civic habits.

Taken together, this empirical framework tests whether voters respond to administrative failure not only at the intensive margin—by changing whether they vote—but also at the extensive margin of vote mode, and whether such responses are transitory or durable. The next section presents the results of this analysis and explores the heterogeneity of these effects by voting method and political party.

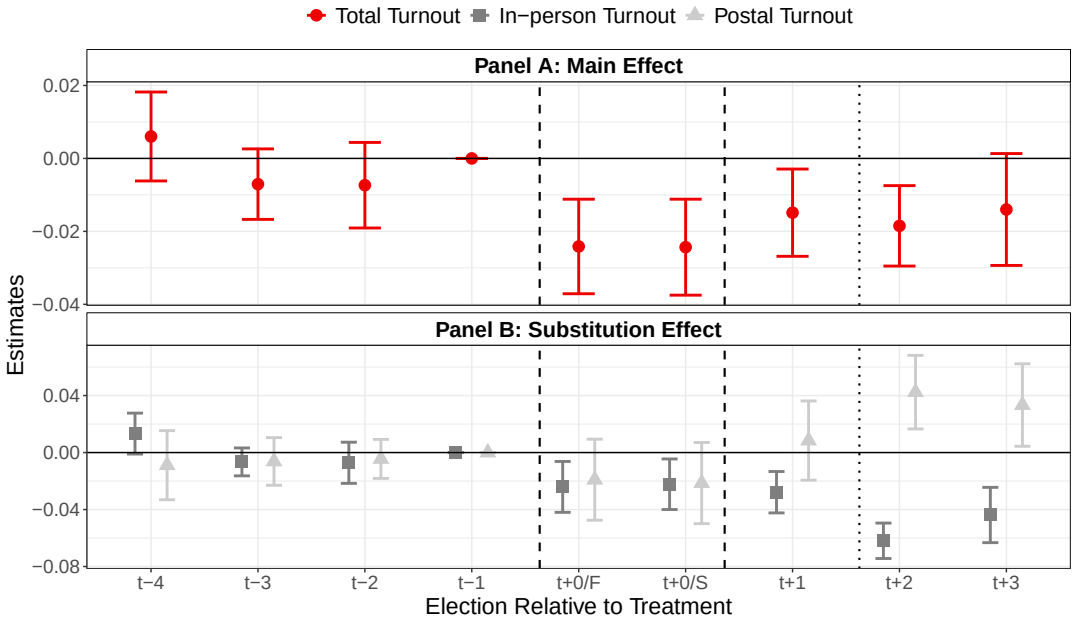
4 Results

4.1 Event Study Results

This section presents the main results from the event study design, which estimates the causal effect of administrative irregularities during the 2021 Berlin State and Federal Elections on voter turnout. Using Equation 5, I track changes in total turnout, in-person voting, and postal voting across treated and untreated precincts over multiple election cycles, interpreting each election as a relative period $\tau \in \{-4, -3, -2, -1, 0, +1, +2, +3\}$ with respect to the 2021 disruption. Figure 3 visualizes the estimated coefficients $\hat{\beta}^{\tau}$ by mode of voting, with the 2019 European election as the reference period $(t - 1)$. The underlying point estimates and standard errors are available in Appendix Table A6.

[3] $0/F$ and $0/S$ refer to the 2021 Federal Election and the 2021 State Election, which took place at the same time.

Figure 3: Main Specification



Notes: The figure presents event study results based on Equation 5 for the natural log of postal, in-person, and total turnout. The reference election $(t-1)$ is the 2019 European election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level. The point estimates and standard errors underlying the results appear in Table A6.

The pre-treatment period serves to assess the plausibility of the parallel trends assumption underlying the identification strategy. Estimated effects for total and in-person turnout in $t-2$ and $t-3$ are both small and statistically indistinguishable from zero, bolstering confidence in the empirical design. A marginally significant increase in postal turnout in $t-3$ (approximately 0.34 percentage points, $p < 0.10$) is noted, though the absence of persistent pre-trends mitigates concerns of confounding anticipation effects.

The 2021 Federal election $(t+0)$, coinciding with the administrative failures, shows a significant drop in total turnout of 1.39 percentage points ($p < 0.01$). This decline is almost entirely attributable to a contraction in in-person voting, which fell by 1.12 percentage points ($p < 0.01$). Postal turnout decreased marginally by 0.26 percentage points but this estimate is not statistically significant ($p = 0.27$), consistent with the institutional constraint that prohibits switching to vote-by-mail on election day. A similar pattern emerges in the simultaneously held 2021 State election, where total turnout fell by 1.44 percentage points, in-person turnout by 1.05 percentage points, and postal turnout remained statistically unchanged.

In the 2023 State election $(t+1)$—a re-run mandated by the Berlin Constitutional Court—total turnout remains depressed by 0.83 percentage points relative to the 2019 baseline. This decline is again driven by persistently lower in-person participation (down 0.72 percentage points), while postal turnout rises slightly by 0.34 percentage points. Although this compensatory increase in mail voting is statistically insignificant, it hints at a gradual behavioral adjustment in the wake of the disruption.

By the time of the 2024 European election ($t + 2$), a more pronounced shift in voting behavior emerges. In-person turnout in treated precincts is 2.32 percentage points lower than in 2019 ($p < 0.01$), marking the largest estimated effect across all post-treatment periods. This drop is partially offset by a statistically significant increase in postal voting of 1.17 percentage points ($p = 0.04$), resulting in a net reduction in total turnout of 1.14 percentage points ($p < 0.05$). These findings suggest that while some voters eventually transition to voting by mail, the adjustment is incomplete and attenuated over time.

Taken together, the estimates provide compelling evidence that administrative irregularities exert both immediate and persistent effects on electoral participation. The decline in in-person voting is sharpest at the moment of disruption, consistent with the theoretical mechanism of increased voting costs suppressing turnout (e.g., Downs, 1957; Pettigrew, 2021). That these effects persist into subsequent elections—even those without observable irregularities—points to behavioral scarring or institutional distrust. Voters exposed to chaos or long queues may revise their beliefs about the costs and reliability of the electoral process, updating their expected utility from voting downward.

The eventual uptick in postal turnout by $t + 2$ aligns with a substitution mechanism: some affected voters appear to adapt their behavior in favor of mail voting. However, this shift does not fully close the participation gap, which remains significant three years after the initial disruption. This echoes evidence from other settings in which negative experiences with electoral administration produce durable declines in civic engagement (Cantoni, 2020; Fujiwara, 2015).

I conduct a range of robustness checks, including alternative fixed effect structures, placebo treatments, and subsample restrictions. These are detailed in Appendix A and Appendix B. Across specifications, the estimated treatment effects on in-person turnout remain statistically significant and of similar magnitude, reinforcing the conclusion that administrative failures have long-lasting consequences for democratic participation.

4.2 Extension: Wait Times

While the main analysis relies on a court-validated binary treatment indicator to capture exposure to electoral irregularities, not all treated precincts experienced disruptions of equal severity. As a complementary extension, this section investigates whether the intensity of disruption—measured by average queue length at peak voting hours—is associated with differential effects on turnout. In doing so, I treat administrative failure as a continuous rather than dichotomous phenomenon, and assess how longer wait times on election day shape subsequent participation.

To operationalize disruption intensity, I use precinct-level estimates of the maximum queue wait time per postal precinct, as given by the *Berlin Constitutional Court ruling*, in hours.

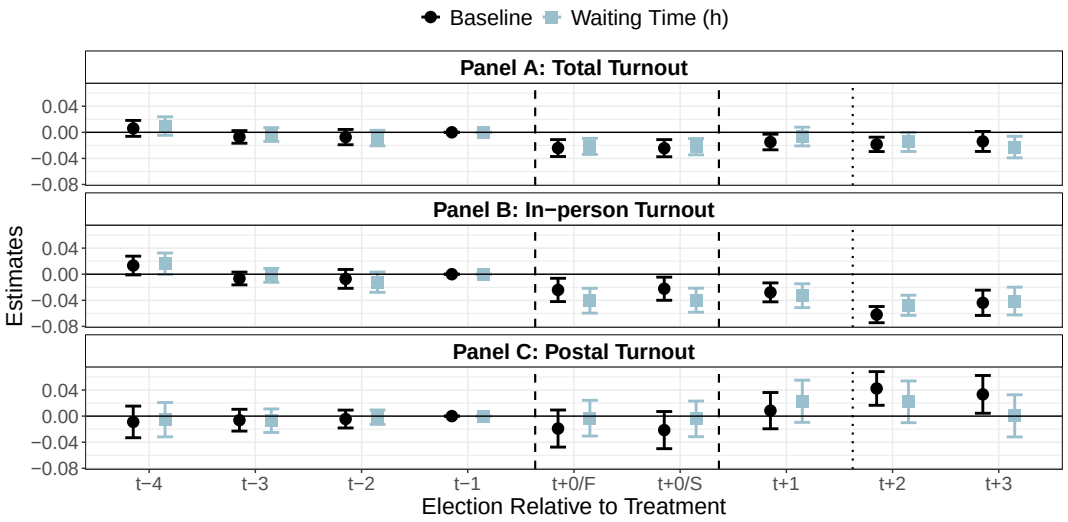
Figure 4 reports the results. In-person turnout declines sharply and significantly in response to longer wait times during the disrupted 2021 Federal and State elections: an additional hour of queuing is associated with a 2.6–2.7% decline in in-person turnout for those contests. The magnitude of this effect increases over time, reaching over 4% in the 2024 European Parliament election, suggesting persistent behavioral disengagement. Postal turnout rises slightly in

response, with increases of 1.8–3.0% in later elections, partially offsetting the drop in in-person participation.

Total turnout effects remain negative and statistically significant, though attenuated relative to in-person results. For example, an additional hour of queuing corresponds to a 1.36 percentage point drop in total turnout in BT21 and AH21. These estimates mirror those found in the binary treatment specification, lending credibility to the interpretation that administrative failure—whether defined legally or operationally—induces lasting reductions in democratic participation.

These results underscore the role of service quality in shaping electoral behavior. Long queues are not merely logistical inconveniences; they represent a signal of bureaucratic inefficiency and procedural unfairness, particularly salient for politically vulnerable populations. The dose-response nature of the effects further supports the interpretation that citizens respond proportionally to the severity of disruption. From a policy perspective, the findings highlight the need to monitor and reduce wait times not only as a matter of convenience, but as a safeguard of electoral inclusion and democratic legitimacy.

Figure 4: Treatment Intensity



Notes: The figure presents triple-interaction estimates based on Equation 5 total turnout (0-100) with a continuous treatment (waiting time in hours). The sample is restricted to postal precincts that are affected by longer waiting times, disruptions, or longer opening hours, and the control group. The reference election ($t-1$) is the 2019 European Election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level. The point estimates and standard errors underlying the results appear in Appendix Table A7. All first and second-order interaction terms required for the identification of the triple-difference estimator are included in the specification or absorbed by the fixed effects.

4.3 Heterogeneities

While the baseline estimates provide the average treatment effect of electoral irregularities on voter turnout, it is likely that such disruptions disproportionately affect certain socio-demographic groups. This section explores heterogeneous effects across precincts by interacting the treatment indicator with standardized socio-demographic characteristics in a triple-difference framework:

$$Y_{idfst} = \sum_{\tau \neq -1} \gamma^{\tau}(\mathbb{1}_{\tau=t} \times treatment_i) + \sum_{\tau \neq -1} \mu^{\tau}(\mathbb{1}_{\tau=t} \times treatment_i \times Z_i)\eta_i + \eta_{dt} + \eta_{ft} + \eta_{st} + \epsilon_{idfst}. \tag{6}$$

Here, Z_i denotes the standardized value of the socio-demographic variable of interest. The coefficient γ^{τ} captures the differential treatment effect at relative time τ, depending on the value of Z_i.

Migrant Background Turnout losses are systematically larger in precincts with a higher share of adult citizens with a migration background. A one standard deviation increase in the migrant share is associated with an additional decline of approximately 0.5 percentage points in total turnout at $t + 0$. This differential persists over time, indicating more sustained disengagement among these voters even beyond the immediate electoral cycle.

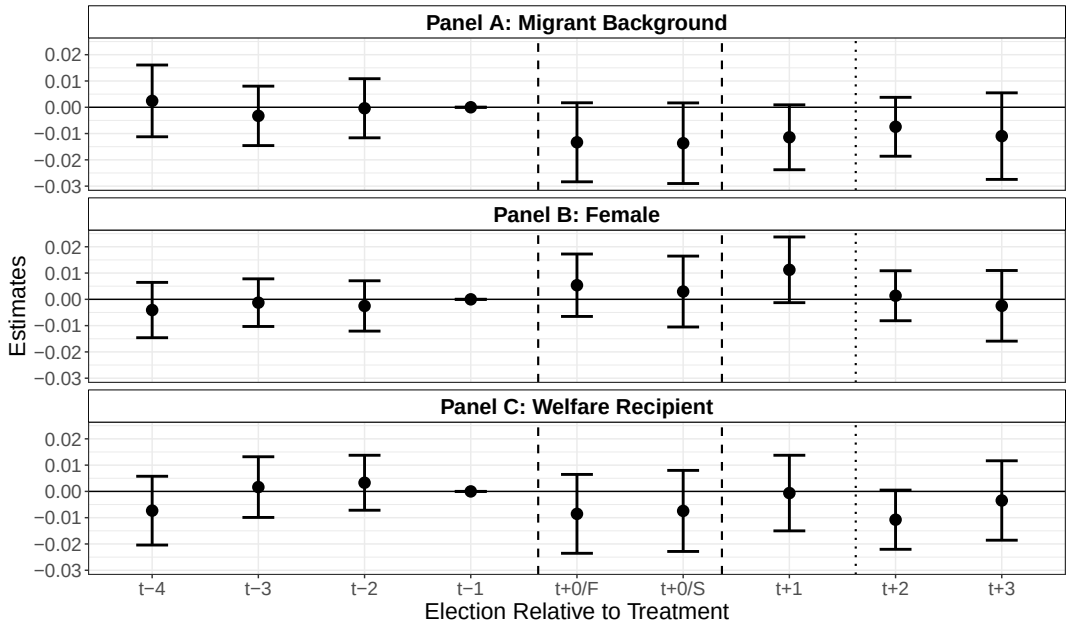
Gender Across election cycles, I find no consistent evidence that treatment effects vary significantly by the share of female voters in a precinct. The interaction coefficients remain statistically indistinct from zero throughout the event window, suggesting that gender composition does not systematically shape the electoral response to irregularities.

Welfare Recipients In contrast, precincts with higher concentrations of welfare recipients exhibit more pronounced and persistent treatment effects. The interaction becomes statistically significant beginning in the first post-treatment election $(t + 1)$ and peaks in $t + 2$, suggesting delayed but stronger disengagement dynamics. This may reflect a compounding of institutional mistrust or perceived marginalization among more economically vulnerable groups.

Age Gradient To examine whether electoral disruptions differentially affect younger versus older voters, I disaggregate by age groups. Results indicate a pronounced **age gradient**: younger voters (aged 18–25 and 25–35) experience the largest negative turnout effects immediately after the treatment, whereas older voters (60+) show no turnout decline and even slight positive deviations in later elections. These patterns suggest that older voters may possess greater institutional resilience or voting habit strength, buffering them against procedural failures. Conversely, the participation of younger cohorts appears more fragile and susceptible to disillusionment. Figure 6 shows the results.

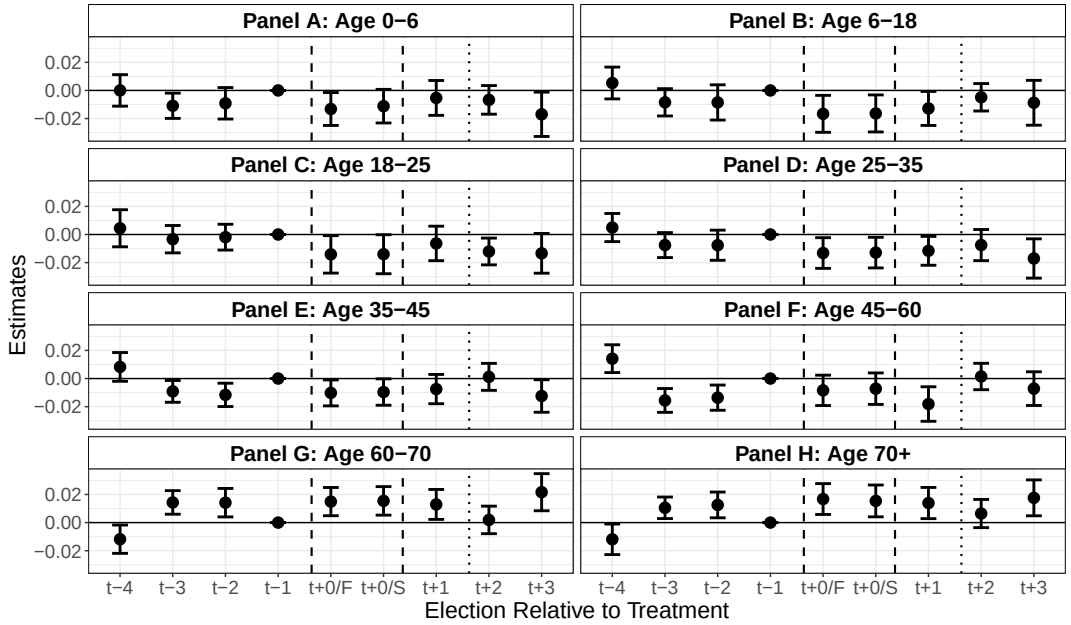
Type of Irregularity While the core empirical design aggregates disruptions into a single treatment indicator, election-day irregularities varied substantially in nature and timing. In particular, disruptions occurred during either the Federal or the State election held concurrently in 2021, allowing for an indirect exploration of heterogeneity by election tier—even though many irregularities overlapped and resist clean categorization.

Figure 5: Heterogeneity by Socio-Demographics



Notes: The figure presents triple-interaction estimates based on Equation 6 total turnout (0-100). The reference election $(t-1)$ is the 2019 European Election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level. The point estimates and standard errors underlying the results appear in Appendix Table A8. All first and second-order interaction terms required for the identification of the triple-difference estimator are included in the specification or absorbed by the fixed effects.

Figure 6: Heterogeneity by Age



Notes: The figure presents triple-interaction estimates based on Equation 6 total turnout (0-100). The reference election $(t-1)$ is the 2019 European Election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level. The point estimates and standard errors underlying the results appear in Appendix Table A9. All first and second-order interaction terms required for the identification of the triple-difference estimator are included in the specification or absorbed by the fixed effects.

Table 1: Treatment Overview

Treatment Federal	Treatment State	Ballot Error	Ballot Missing	Disruption	Unusual Waiting	Open After Official Closing	N
yes	*yes*	32	13	101	90	212	277
no	*yes*	45	61	2	55	1	157
Subsample: Disruption = 0 & Open A.O. Closing = 0							
yes	*yes*	1	2	–	16	–	17
no	*yes*	44	61	–	55	–	154

Notes: The table presents the number of precincts by treatment during the Federal and State elections. The type of irregularity is available only for the State election.

The most severe disruptions coincided with the Federal election and involved visible operational failures: polling stations accepting votes after the legal closing time, redirecting voters between precincts, and remaining open well past 6:00 p.m. By contrast, State-level irregularities more often involved logistical mishaps such as incorrect or missing ballots. These may have been less visible to voters or more easily corrected before deterring participation.

To probe whether the timing and salience of disruptions matter for electoral behavior, I estimate separate event-study specifications for three mutually exclusive groups: (i) precincts affected during the Federal election (my primary treatment group), (ii) precincts affected only during the State election, and (iii) unaffected precincts, which serve as the counterfactual.

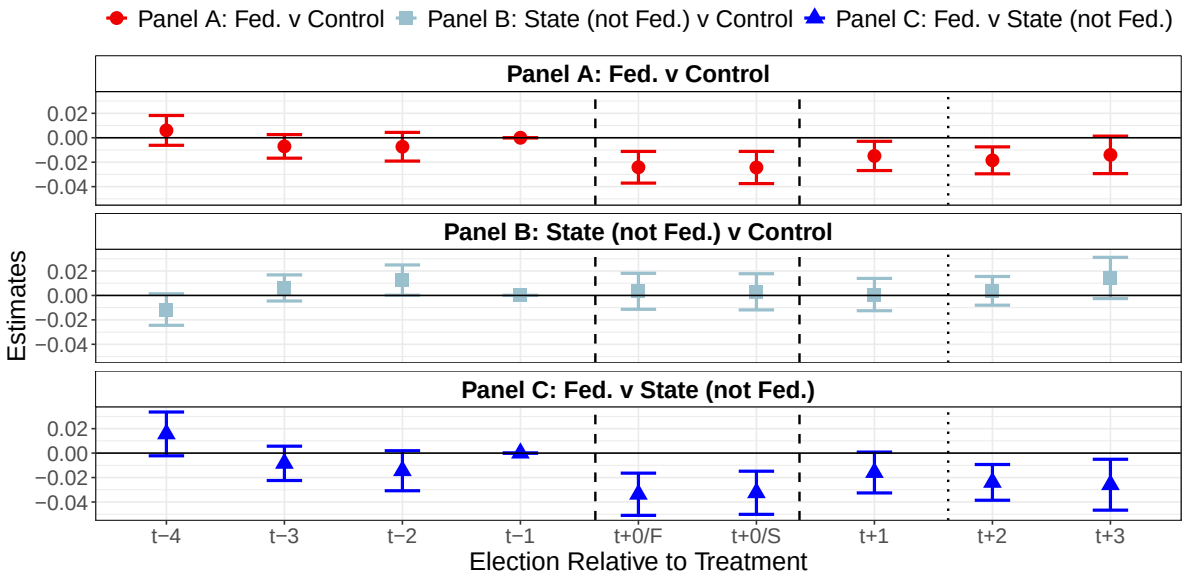
Importantly, these estimates are not based on a triple-difference design. Instead, I disaggregate treatment timing and compare turnout trajectories across the three groups.

Figure 7 presents the resulting trends. Turnout suppression is concentrated in precincts disrupted during the Federal election: the immediate decline at $t{+}0$ is statistically significant and remains directionally negative in subsequent elections. By contrast, precincts with State-only disruptions show no detectable turnout effects, suggesting that these incidents lacked sufficient salience or severity to shift electoral engagement.

This interpretation is supported by descriptive evidence on the incidence of specific irregularities. Table 1 tabulates the frequency of ballot-related errors, operational disruptions, and violations of closing time rules. Notably, precincts affected during the Federal election were disproportionately more likely to experience highly salient procedural breakdowns—including 212 cases of late closing and 101 instances classified as general disruptions. In contrast, precincts affected only at the State level saw virtually no such violations.

Taken together, these findings indicate that voter disengagement is not simply a function of treatment status, but emerges particularly when disruptions undermine the visible legitimacy or procedural reliability of the electoral process.

Figure 7: Heterogeneity by Type of Treatment



Notes: The figure presents triple-interaction estimates for postal turnout, in-person turnout, and total turnout (0-100). The reference election ($t-1$) is the 2019 European Election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level.

5 Attitudinal Evidence from Survey Data

To test whether the 2021 administrative disruptions affected voters' beliefs and perceptions, I complement the turnout analysis with survey evidence from the German Longitudinal Election Study (GLES). I use repeated rolling cross-sections from the 2017, 2021, and 2025 federal elections and estimate a difference-in-differences specification comparing Berlin residents to respondents in other states. The estimation follows Equation 7.

The first outcome captures trust in electoral integrity—operationalized as agreement with the statement. This outcome directly maps to the informational updating mechanism described in subsection 3.1.

$$Y_{ist} = \sum_{\kappa \neq 2017} \beta^{\kappa}(\mathbb{1}_{\kappa=t} \times \text{Berlin}_i) + \gamma_s + \lambda_t + u_{ist}, \tag{7}$$

where Y_{ist} is the outcome for individual i in state s and year t, with state (γ_s) and year (λ_t) fixed effects. Standard errors are clustered at the state level. The omitted year is 2017.

Panel A of Table 2 shows a significant decline in trust in Berlin in 2021, with effects that attenuate by 2025. This supports the interpretation that salient administrative failures produced a short-term revision in beliefs about institutional competence—consistent with the informational updating channel in the model.

Panel B investigates the timing of abstention decisions. Estimates show that nonvoters in Berlin were substantially more likely to report late-stage abstention in 2021. Though based on a smaller sample ($n \approx 510$), the effects are large and statistically significant.

Together, these findings provide survey-based evidence consistent with the model in subsection 3.1. The observed trust decline reflects belief updating after visible state failure, while the shift in abstention timing suggests heightened perceived costs. Although institutional trust appears to rebound by 2025, turnout remains persistently lower, suggesting that even short-lived informational shocks may induce lasting behavioral disengagement via disrupted civic routines.

Table 2: Robustness: Controls

Panel A: "Process Correct & Fair"	(1)	(2)	(3)
Berlin (t+0)	-0.0876***	-0.1379***	-0.5935***
	(0.0038)	(0.0050)	(0.0224)
Berlin (t+1)	-0.0122***	-0.0008	-0.0413*
	(0.0031)	(0.0035)	(0.0213)
Observations	10,443	14,700	16,310
Specification	strongly (dis)agree	(strongly) (dis)agree	continuos
Year FE	✓	✓	✓
State FE	✓	✓	✓
Panel B: Timing Decision Abstain	**(1)**	**(2)**	
Berlin (t+0)	0.4950***	0.5281***	
	(0.0401)	(0.0360)	
Berlin (t+1)	0.1823**	-0.0077	
	(0.0621)	(0.0527)	
Observations	510	510	
Specification	Same Day + Prevented	Prevented	
Year FE	✓	✓	
State FE	✓	✓	

Notes: This table reports estimates from Equation 7 using pooled data from the GLES Rolling Cross-Section surveys conducted in 2017, 2021, and 2025. *Panel A* reports respondents' beliefs about electoral integrity, measured via agreement with the statement: *"Die Bundestagswahl wurde von den zuständigen Behörden korrekt und fair durchgeführt"*. Binary dependent variables equal 1 if respondents "fully agree" (Column 1), or "fully" or "somewhat agree" (Column 2); Column (3) codes responses from -2 ("strongly disagree") to $+2$ ("strongly agree"). *Panel B* examines the timing of abstention decisions among nonvoters. Column (1) codes as 1 those who decided not to vote in the final days before the election or reported being prevented from voting. Column (2) isolates respondents who cited inability to vote as the main reason. The reference election ($t-1$) is the 2017 Federal Election. Standard errors are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6 Downstream Electoral Consequences

Election-day disruptions may not only reduce aggregate turnout but also influence the political composition of the electorate. This section investigates whether precinct-level irregularities during the 2021 elections had measurable effects on the vote shares of major political parties.

I estimate a series of stacked event study models using the log vote share of individual parties as the dependent variable. This specification captures proportional changes in party support over time and facilitates comparability across parties with different baseline levels. The

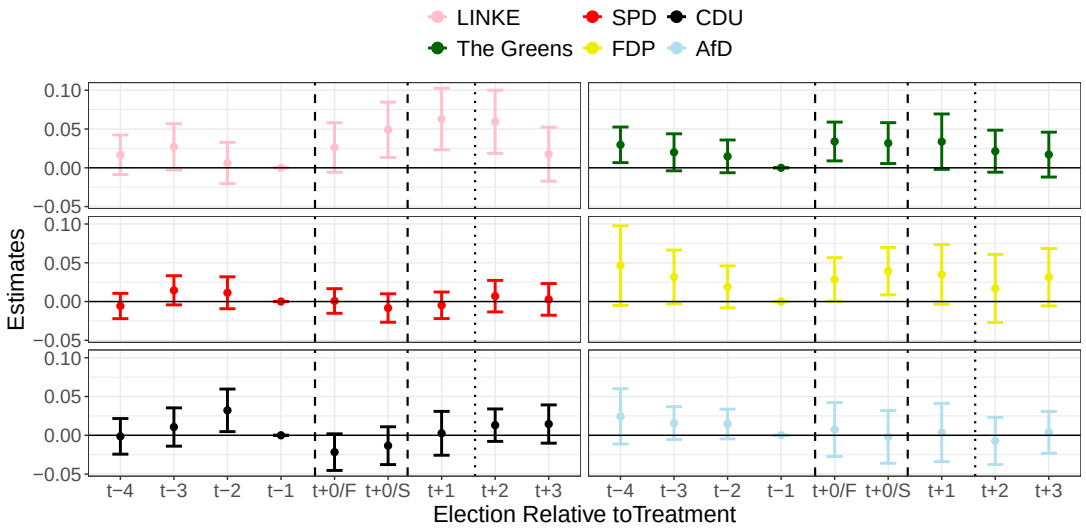
reference election is the 2019 European Parliament election, and all estimates include fixed effects for precincts, election districts, and federal/state contexts.

?? presents the estimated effects for all major parties from four elections prior to three elections after the treatment. Pre-treatment trends are flat and statistically indistinguishable from zero, supporting the credibility of the identification strategy.

Following the 2021 disruptions ($t = 0$), there is no immediate or statistically significant shift in support for any major party. In subsequent elections, two parties display measurable changes: the FDP exhibits a small but statistically significant increase in vote share, while the CDU shows a significant decline. These patterns emerge at $t + 1$ and persist into $t + 2$, suggesting asymmetric effects in electoral engagement or mobilization following the disruptions. For all other parties—including SPD, Greens, the Left (LINKE), and AfD—vote share changes remain statistically indistinct from zero throughout the post-treatment period.

Taken together, these results indicate that administrative irregularities can generate modest but detectable shifts in electoral outcomes at the party level, even absent changes in overall turnout. While most effects are null, the divergence between FDP and CDU responses underscores that disruptions can produce distributive consequences among opposition parties, possibly due to differences in voter profile, campaign mobilization, or susceptibility to procedural legitimacy concerns.

Figure 8: Election Results (Vote Share)



Notes: The figure presents event study results based on Equation 5 for the total vote share of the governing party at the time of the treatment – Social Democrats, the Green Party, and the Left –, the opposition – the conservatives, the Liberals, and a right-wing populist party –, and the party *Alternative für Deutschland* (AfD), a right-wing populist party. The reference election ($t-1$) is the 2019 European election. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the postal precinct level. The point estimates and standard errors underlying the results appear in Appendix Table A10.

7 Conclusion

This paper demonstrates that even in high-capacity democratic settings, administrative failures in election logistics can meaningfully distort political participation. Exploiting quasi-experimental variation from Berlin's 2021 electoral crisis—when severe irregularities disrupted voting procedures in hundreds of precincts—I provide causal evidence that these disruptions depressed turnout by approximately 1.2 percentage points, with effects persisting across multiple electoral cycles. Voters in affected precincts were not only discouraged from participating on election day, but many remained disengaged years later.

The behavioral response is concentrated among groups already at risk of political marginalization. Voters with migration backgrounds, younger cohorts, and welfare recipients exhibit larger and more persistent reductions in turnout. These patterns are consistent with theories of administrative burden and habit formation, where early negative experiences with the state's procedural reliability can scar civic engagement.

While aggregate turnout falls, political consequences remain muted. Most parties do not experience significant shifts in vote share, suggesting that voters do not simply switch allegiance but are more likely to abstain. The only detectable changes are asymmetric and modest: the FDP sees a slight gain, while the CDU loses support. These results suggest that administrative failures may have distributive effects, but they are contingent on party-specific mobilization or voter composition, not systematic partisan bias.

From a policy perspective, the findings underscore that electoral integrity extends beyond ballot security or fraud prevention—it includes the routine, mundane, but essential task of competent election logistics. Failures in these domains can suppress participation, deepen representational gaps, and damage institutional trust. Improving electoral resilience—through contingency planning, better staffing, and clearer accountability—should thus be viewed as a democratic investment.

More broadly, this study highlights that elections are not merely moments of preference aggregation but instances of public service delivery. When that service breaks down, especially in visible and salient ways, it sends a signal about state competence and responsiveness. Voters, particularly those on the margins of political inclusion, receive and respond to that signal. As governments grapple with declining turnout and rising polarization, ensuring reliable and equitable access to the ballot box remains foundational to democratic legitimacy.

References

Akhtari, M., Moreira, D., and Trucco, L. (2022). Political turnover, bureaucratic turnover, and the quality of public services. *Quarterly Journal of Economics*, 137(4):2385–2432.

Alipour, J.-V. and Lindlacher, V. (2025). No surprises, please: Voting costs and electoral turnout. *Journal of Political Economy Microeconomics*, 3(1):000–000.

Alsan, M. and Wanamaker, M. (2018). Tuskegee and the health of black men. *The quarterly journal of economics*, 133(1):407–455.

Bechtel, M. M., Hangartner, D., and Schmid, L. (2018). Compulsory voting, habit formation, and political participation. *Review of Economics and Statistics*, 100(3):467–476.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics*, 12(1):61–85.

Cantoni, E. and Pons, V. (2021). Strict id laws don't stop voters: Evidence from a us nationwide panel, 2008–2018. *The Quarterly Journal of Economics*, 136(4):2615–2660.

Cantoni, E., Pons, V., and Schafer, J. (2025). Voting rules, turnout, and economic policies. *Annual Review of Economics*, 17.

Downs, A. (1957). An economic theory of democracy. *Harper and Row*, 28.

Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464.

Fujiwara, T., Meng, K., and Vogl, T. (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics*, 8(4):160–188.

Gaebler, S., Potrafke, N., and Roesel, F. (2020). Compulsory voting and political participation: Empirical evidence from austria. *Regional Science and Urban Economics*, 81:103499.

Gerber, A. S., Green, D. P., and Shachar, R. (2003). Voting may be habit-forming: evidence from a randomized field experiment. *American journal of political science*, 47(3):540–550.

Gottlieb, J. (2016). Greater expectations: A field experiment to improve accountability in mali. *American Journal of Political Science*, 60(1):143–157.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20(1):25–46.

Heß, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal*, 17(3):630–651.

Hoffman, M., León, G., and Lombardi, M. (2017). Compulsory voting, turnout, and government spending: Evidence from austria. *Journal of Public Economics*, 145:103–115.

Kaplan, E. and Yuan, H. (2020). Early voting laws, voter turnout, and partisan vote composition: Evidence from ohio. *American Economic Journal: Applied Economics*, 12(1):32–60.

Lowes, S. and Montero, E. (2021). The legacy of colonial medicine in central africa. *American Economic Review*, 111(4):1284–1314.

Pettigrew, S. (2021). The downstream consequences of long waits: How lines at the precinct depress future turnout. *Electoral studies*, 71:102188.

Potrafke, N. and Roesel, F. (2020). Opening hours of polling stations and voter turnout: Evidence from a natural experiment. *The Review of International Organizations*, 15(1):133–163.

Riker, W. H. and Ordeshook, P. C. (1968). A theory of the calculus of voting. *American political science review*, 62(1):25–42.

Rosas, G. (2010). Trust in elections and the institutional design of electoral authorities: Evidence from latin america. *Electoral Studies*, 29(1):74–90.

Toral, G. (2023). How patronage delivers: Political appointments, bureaucratic accountability, and service delivery in brazil. *American Journal of Political Science*.

Online Appendix

Race to the Polls: The Berlin Marathon and its Electoral Consequences
by *Marius Kröper*

A Placebo: Randomization Inference

I utilize a randomization inference approach following Heß (2017), conducting 999 permutations while maintaining the number of treated municipalities within each district. This method assesses whether the observed results could arise by random chance. The results, presented in Table A1, show significant treatment effects consistent with the original analysis, thus reinforcing the reliability of my findings. This approach helps to ensure that the results are due to the actual treatment effect of the irregularities during the 2021 Federal elections and not random variation.

Table A1: Randomized Inference

	Total Turnout (1)	In-person Turnout (2)	Postal Turnout (3)
Treatement (t-4)	0.0138	0.3645	-0.3507
	(0.2945)	(0.2387)	(0.2275)
	[0.9710]	[0.1021]	[0.1151]
Treatement (t-3)	-0.4513	-0.1428	-0.3085*
	(0.2826)	(0.2038)	(0.1677)
	[0.1502]	[0.5105]	[0.0641]*
Treatement (t-2)	-0.3090	-0.2065	-0.1025
	(0.3131)	(0.3027)	(0.1634)
	[0.3333]	[0.5125]	[0.5726]
Treatement (t+0/Federal)	-1.390***	-1.122***	-0.2684
	(0.3503)	(0.3873)	(0.4113)
	[0.0000]***	[0.0040]***	[0.5205]
Treatement (t+0/State)	-1.437***	-1.047***	-0.3899
	(0.3671)	(0.3861)	(0.4172)
	[0.0000]***	[0.0100]**	[0.3413]
Treatement (t+1)	-0.8260**	-1.162***	0.3361
	(0.3884)	(0.2949)	(0.3725)
	[0.0430]**	[0.0000]***	[0.4154]
Treatement (t+2)	-1.143***	-2.317***	1.173***
	(0.3778)	(0.2601)	(0.3695)
	[0.0040]***	[0.0000]***	[0.0040]***
R^2	0.95213	0.89287	0.92054
Observations	9,808	9,808	9,808
Precinct FE	✓	✓	✓
Election-District FE	✓	✓	✓

Notes: The table presents event study results based on Equation 5 for postal turnout, in-person turnout, and total turnout (0-100). The p-values calculated based on the randomized inference approach following Heß (2017) are shown in brackets. The reference election $(t-1)$ is the 2019 European Election. Standard errors are clustered at the postal precinct level and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Further Robustness

Controls In the main specification, I only control for the log number of eligible voters held constant to the 2017 Federal Election, the last federal election before the treatment, and interact them with the election indicators. To test the robustness of the main specification, I add further time-invariant controls, holding them constant for 2019. Controls are the share of the following groups: German citizens with migrant backgrounds, foreigners, persons subject to social security, 6-18 year old, 18-25 year old, 65 to 99 year old, EU citizens, persons living in a civil union, Single households; additionally I control for the log number of eligible voters. Results are shown in Table A2

Table A2: Robustness: Controls

	Total Turnout	In-person Turnout	Postal Turnout
	(1)	(2)	(3)
Treatement (t-4)	0.0055	0.0129**	-0.0098
	(0.0058)	(0.0065)	(0.0122)
Treatement (t-3)	-0.0049	-0.0043	-0.0040
	(0.0041)	(0.0043)	(0.0080)
Treatement (t-2)	-0.0069	-0.0061	-0.0061
	(0.0043)	(0.0052)	(0.0066)
Treatement (t+0/Federal)	-0.0220***	-0.0252***	-0.0181
	(0.0063)	(0.0071)	(0.0141)
Treatement (t+0/State)	-0.0221***	-0.0234***	-0.0201
	(0.0065)	(0.0071)	(0.0140)
Treatement (t+1)	-0.0151***	-0.0288***	0.0049
	(0.0057)	(0.0076)	(0.0137)
Treatement (t+2)	-0.0231***	-0.0618***	0.0284**
	(0.0055)	(0.0066)	(0.0132)
Treatement (t+3)	-0.0193***	-0.0425***	0.0157
	(0.0063)	(0.0075)	(0.0142)
R^2	0.94967	0.91415	0.93792
Observations	12,150	12,150	12,150
Precinct FE	✓	✓	✓
Election-District FE	✓	✓	✓

Notes: The table presents event study results based on Equation 5 for postal turnout, in-person turnout, and total turnout (0-100) controlling for additional variables. The reference election ($t-1$) is the 2019 European Election. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Clustering In my baseline analysis, I cluster at the level of postal precincts, which is the level at which the treatment is applied. The district is the next higher administrative unit, with precinct administrators reporting to district offices. Concerns may arise that model errors are correlated within State constituencies. To address this, I increase the clustering level to State constituencies, as shown in Column 2 of Table A3. The standard errors are only marginally larger when clustering at the county level. Additionally, two-way clustering—considering postal precincts, and district × elections results in slightly smaller standard errors (Column 3) compared to clustering at the postal precinct level.

Table A3: Robustness: Standard Errors

Panel A: Total Turnout	(1)	(2)	(3)
Treatement (t-4)	0.0060	0.0060	0.0060
	(0.0062)	(0.0085)	(0.0088)
Treatement (t-3)	-0.0071	-0.0071	-0.0071
	(0.0049)	(0.0054)	(0.0065)
Treatement (t-2)	-0.0073	-0.0073	-0.0073
	(0.0060)	(0.0063)	(0.0072)
Treatement (t+0/Federal)	-0.0241***	-0.0241***	-0.0241***
	(0.0066)	(0.0056)	(0.0066)
Treatement (t+0/State)	-0.0243***	-0.0243***	-0.0243***
	(0.0067)	(0.0057)	(0.0067)
Treatement (t+1)	-0.0149**	-0.0149**	-0.0149*
	(0.0061)	(0.0060)	(0.0077)
Treatement (t+2)	-0.0185***	-0.0185***	-0.0185***
	(0.0056)	(0.0063)	(0.0060)
Treatement (t+3)	-0.0140*	-0.0140**	-0.0140
	(0.0078)	(0.0068)	(0.0096)
Standard-Errors	Precinct	Precinct-Election	Precinct & Election-District
R^2	0.92949	0.92949	0.92949
Panel B: In-person Turnout	**(1)**	**(2)**	**(3)**
Treatement (t-4)	0.0133*	0.0142**	0.0142*
	(0.0073)	(0.0072)	(0.0072)
Treatement (t-3)	-0.0066	-0.0071	-0.0071
	(0.0050)	(0.0049)	(0.0049)
Treatement (t-2)	-0.0072	-0.0081	-0.0081
	(0.0074)	(0.0072)	(0.0072)
Treatement (t+0/Federal)	-0.0241***	-0.0256***	-0.0256***
	(0.0091)	(0.0089)	(0.0089)
Treatement (t+0/State)	-0.0222**	-0.0237***	-0.0237***
	(0.0090)	(0.0088)	(0.0088)
Treatement (t+1)	-0.0278***	-0.0284***	-0.0284***
	(0.0074)	(0.0074)	(0.0074)
Treatement (t+2)	-0.0619***	-0.0619***	-0.0619***
	(0.0063)	(0.0064)	(0.0064)
Treatement (t+3)	-0.0438***	-0.0446***	-0.0446***
	(0.0099)	(0.0097)	(0.0098)
R^2	0.86623	0.86886	0.86886
Panel C: Postal Turnout	**(1)**	**(2)**	**(3)**
Treatement (t-4)	-0.0089	-0.0091	-0.0091
	(0.0124)	(0.0124)	(0.0124)
Treatement (t-3)	-0.0062	-0.0069	-0.0069
	(0.0085)	(0.0085)	(0.0085)
Treatement (t-2)	-0.0044	-0.0049	-0.0049
	(0.0070)	(0.0069)	(0.0070)
Treatement (t+0/Federal)	-0.0191	-0.0193	-0.0193
	(0.0145)	(0.0145)	(0.0145)
Treatement (t+0/State)	-0.0214	-0.0216	-0.0216
	(0.0145)	(0.0145)	(0.0145)
Treatement (t+1)	0.0084	0.0078	0.0078
	(0.0142)	(0.0142)	(0.0142)
Treatement (t+2)	0.0424***	0.0416***	0.0416***
	(0.0132)	(0.0131)	(0.0132)
Treatement (t+3)	0.0333**	0.0321**	0.0321**
	(0.0148)	(0.0146)	(0.0147)
R^2	0.93030	0.93054	0.93054
Observations	12,150	12,150	12,150
Precinct FE	✓	✓	✓
Election-District FE	✓		✓
Election-District-Federal const. FE		✓	
Election-Federal const. FE			✓

Notes: The table presents event study results based on Equation 5 for postal turnout, in-person turnout, and total turnout (0-100) using different standard error calculation methods: Column (1) is the baseline and standard errors are clustered at the municipality level. In Column (2), standard errors are clustered at the State conctituency level. Column (3) uses two-way clustered standard errors at the level of postal precinct and District × elections. Standard errors are reported in parentheses. The reference election $(t-1)$ is the 2019 European Election. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Matching on Observables I conduct various matching procedures to ensure the comparability of the treatment and control group, namely propensity score matching, Entropy Balancing as proposed by Hainmueller (2012), and Mahalanobis distance matching. If applicable, I solely refer to values from the year 2017, as this is when the last pre-treatment federal election took place. I re-evaluate my baseline model (Equation 5) using each matching method and present the findings in Appendix Table A4.

Propensity Score Matching I conduct a propensity score matching procedure to estimate the likelihood of experiencing irregularities during the 2021 Federal elections. This propensity is calculated using a probit regression based on the following pre-treatment socioeconomic characteristics: population measured by the number of eligible voters, share of population aged 6-18, 18-25, and over 65, share of foreign population, share of EU citizens, share of females, share of citizens with migrant backgrounds, share of single households, and share of residents subject receiving social security. Additionally, I exact match on the district. All precinct characteristics are measured as of 2019. The matched sample is obtained using 1:1 nearest neighbor matching with replacement. In this process, 282 treated units are matched to 153 control units. The estimates derived from the matched sample confirm my original results for the effect of irregularities on total turnout with the exception of the 2023 State election, which loses significance.

Mahalanobis Matching I check the robustness of my findings by matching treated and untreated units based on the Mahalanobis distance. This method evaluates similarity by measuring the proximity of units' covariates in vector space. I calculate distances using the same pre-treatment covariates as previously mentioned and employ a 1:1 nearest neighbor matching with replacement to create the matched sample. In this process, 282 treated municipalities are matched to 151 control units, while 793 units are dropped. The results indicate that the treatment effects remain significant and are consistent with the outcomes of propensity score matching.

Entropy Balancing I implement the entropy balancing approach proposed by Hainmueller (2012). This method offers the advantage of not truncating the sample. Instead, it assigns a set of weights that balance the treatment and control groups across several moments of the covariate distributions. Specifically, I balance the means and variances of the pre-treatment characteristics listed above and use the resulting entropy weights in the event study regressions. The results consistently support my original findings across all specifications and outcomes.

Table A4: Matching

	Total Turnout		
	(1)	(2)	(3)
Treatement (t-4)	0.0051	0.0044	0.0029
	(0.0078)	(0.0080)	(0.0067)
Treatement (t-3)	-0.0065	-0.0045	0.0070
	(0.0061)	(0.0065)	(0.0046)
Treatement (t-2)	-0.0043	-0.0042	0.0093
	(0.0077)	(0.0079)	(0.0059)
Treatement (t+0/Federal)	-0.0200***	-0.0263***	-0.0182***
	(0.0077)	(0.0079)	(0.0059)
Treatement (t+0/State)	-0.0183**	-0.0246***	-0.0179***
	(0.0078)	(0.0080)	(0.0061)
Treatement (t+1)	-0.0152*	-0.0204***	-0.0144**
	(0.0082)	(0.0077)	(0.0064)
Treatement (t+2)	-0.0165**	-0.0206***	-0.0217***
	(0.0079)	(0.0073)	(0.0061)
Treatement (t+3)	-0.0128	-0.0177*	-0.0022
	(0.0101)	(0.0103)	(0.0082)
R^2	0.93037	0.93130	0.92810
Observations	3,942	3,924	11,034
Specification	Propensity Score	Mahalanobis	Entropy Bal.
Precinct FE	✓	✓	✓
Election-District FE	✓	✓	✓

Notes: The table presents event study results based on Equation 5 total turnout (0-100) using different matching approaches. Column (1) uses nearest neighbour matching based on propensity score, and Column (2) based on Mahalanobis distance. Column (3) uses weights from Hainmueller (2012). The reference election ($t-1$) is the 2019 European Election. Standard errors are clustered at the municipality level and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C Data Appendix

Main Administrative and Electoral Data The core dataset consists of precinct-level turnout and party vote shares across all Berlin elections between 2014 and 2025, harmonized via spatial crosswalks. I construct a balanced panel at the level of 2021 postal voting precincts (`BWB21`), the base geography throughout. For each election, I aggregate turnout and valid votes separately for in-person (`UWB`) and postal precincts, reconstructing total electorate size and turnout by mode.

The dataset covers eight elections over six election years: European Parliament elections in 2014, 2019, and 2024; state elections in 2016, 2021, and 2023; and federal elections in 2017 and 2021. Precinct shapefiles for each election year are obtained from the Berlin electoral administration. To link precincts over time, I calculate area-weighted crosswalks using high-resolution 100-meter raster population data from the 2022 census (restricted to German citizens aged 18 and older). This ensures that precinct weights reflect electorally relevant populations. To maintain comparability, I exclude 157 precincts affected only by the state court's ruling.

Electoral results at the precinct level are drawn from official election reports. Data are cleaned, harmonized, and weighted into the 2021 spatial frame using the crosswalks. I construct

turnout and party shares relative to eligible voters, and build separate indicators for postal and in-person turnout. I use postal precincts as the unit of analysis since they report full data on voting mode and vote shares and allow clean assignment of treatment status based on rerun designation.

The analysis also incorporates detailed precinct-level socio-demographics from Berlin's structural 2019 data reports (*Strukturdaten*), including age, gender, migration background, and welfare dependency shares. These are merged via unique precinct identifiers and used in heterogeneity analyses.

Administrative disruption data come from official court decisions on the 2021 Berlin elections, specifying ballot shortages, waiting times, and types of irregularities by polling station. I aggregate these indicators to the `BWB21` level and use them to define treatment and exposure intensities. Treated precincts are those flagged in the Federal Constitutional Court's judgment; court decisions define treatment with legal precision and spatial granularity.

Survey Data: GLES Rolling Cross Sections To complement the administrative data, I draw on the GLES Rolling Cross-Section (RCS) surveys from 2017, 2021, and 2025. These surveys include repeated nationally representative samples of German citizens during each federal election campaign. I use only respondents aged 18 or older with valid information on residence (`State`), trust in election administration, and nonvoting timing.

The relevant trust item asks whether respondents agree with the statement: *"Die Bundestagswahl wurde von den zuständigen Behörden korrekt und fair durchgeführt"*. I construct three variables: (i) a binary indicator for those who ´´fully agree", (ii) a relaxed version including ´´somewhat agree", and (iii) an ordinal scale from -2 ("strongly disagree") to $+2$ (´´strongly agree")).

For nonvoters, the timing of abstention is measured via post-election questions. I define two binary indicators: one for individuals who abstained late (on or shortly before election day), and one for those who cited procedural reasons for not voting. These outcomes allow testing for latent cost shocks, as outlined in the behavioral model.

The estimation strategy compares Berlin respondents to residents in other states, as formalized in Equation 7. Standard errors are clustered at the state level.

D Tables

Table A5: Summary Statistics

Panel A: Full Sample								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
NA	1,350	76.618	9.302	48.053	70.233	77.755	83.153	125.795
NA.1	1,350	39.742	3.604	26.594	37.419	39.783	42.055	56.059
NA.2	1,350	36.876	8.701	13.857	30.586	37.005	42.422	86.934
Eligible Voters	1,350	1,658.284	621.084	24.216	1,161.356	1,417.478	2,158.430	3,912.376
SGCwMB	1,350	2.475	3.108	0.024	0.733	1.494	3.088	35.697
Panel B: Treatment Group								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
NA	294	77.367	7.929	48.053	73.304	78.883	82.978	100.986
NA.1	294	39.868	3.184	31.101	37.839	39.769	41.933	48.584
NA.2	294	37.499	7.433	14.359	33.195	38.568	42.257	62.366
Eligible Voters	294	1,921.515	724.212	651.134	1,245.193	1,908.351	2,604.584	3,912.376
SGCwMB	294	2.833	3.203	0.026	0.878	1.979	3.583	27.875
Panel C: Control Group								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
NA	1,056	76.410	9.643	51.107	69.533	77.041	83.309	125.795
NA.1	1,056	39.707	3.713	26.594	37.363	39.803	42.177	56.059
NA.2	1,056	36.702	9.018	13.857	30.074	36.504	42.458	86.934
Eligible Voters	1,056	1,584.999	568.238	24.216	1,146.627	1,361.326	2,008.343	3,848.979
SGCwMB	1,056	2.375	3.076	0.024	0.687	1.378	2.886	35.697

Notes: The table presents summary statistics for the 2021 Federal election for the main sample, i.e. excluding the district "Tempelhof-Schöneberg" and those municipalities only affected by irregularities in the State election. It is split by treatment and control group for selected variables used in the analyses. *SGCwMB* is the share of German citizens older than 18 with migrant backgrounds.

Table A6: Main Specification

	Total Turnout (1)	In-person Turnout (2)	Postal Turnout (3)
Treatement (t-4)	0.0060	0.0133*	-0.0089
	(0.0062)	(0.0073)	(0.0124)
Treatement (t-3)	-0.0071	-0.0066	-0.0062
	(0.0049)	(0.0050)	(0.0085)
Treatement (t-2)	-0.0073	-0.0072	-0.0044
	(0.0060)	(0.0074)	(0.0070)
Treatement (t+0/Federal)	-0.0241***	-0.0241***	-0.0191
	(0.0066)	(0.0091)	(0.0145)
Treatement (t+0/State)	-0.0243***	-0.0222**	-0.0214
	(0.0067)	(0.0090)	(0.0145)
Treatement (t+1)	-0.0149**	-0.0278***	0.0084
	(0.0061)	(0.0074)	(0.0142)
Treatement (t+2)	-0.0185***	-0.0619***	0.0424***
	(0.0056)	(0.0063)	(0.0132)
Treatement (t+3)	-0.0140*	-0.0438***	0.0333**
	(0.0078)	(0.0099)	(0.0148)
R^2	0.92949	0.86623	0.93030
Observations	12,150	12,150	12,150
Precinct FE	✓	✓	✓
Election-District FE	✓	✓	✓

Notes: The table presents event study results based on Equation 5 for postal turnout, in-person turnout, and total turnout (0-100). The reference election ($t-1$) is the 2019 European Election. Standard errors are clustered at the postal precinct level and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: Main Specification

Panel A: Total Turnout	(1)	(2)
Treatement (t-4)	0.0060	
	(0.0062)	
Treatement (t-3)	-0.0071	
	(0.0049)	
Treatement (t-2)	-0.0073	
	(0.0060)	
Treatement (t+0/Federal)	-0.0241***	
	(0.0066)	
Treatement (t+0/State)	-0.0243***	
	(0.0067)	
Treatement (t+1)	-0.0149**	
	(0.0061)	
Treatement (t+2)	-0.0185***	
	(0.0056)	
Treatement (t+3)	-0.0140*	
	(0.0078)	
Waiting Time (t-4)		0.0096
		(0.0072)
Waiting Time (t-3)		-0.0034
		(0.0054)
Waiting Time (t-2)		-0.0088
		(0.0061)
Waiting Time (t+0/Federal)		-0.0216***
		(0.0063)
Waiting Time (t+0/State)		-0.0221***
		(0.0064)
Waiting Time (t+1)		-0.0065
		(0.0073)
Waiting Time (t+2)		-0.0148**
		(0.0074)
Waiting Time (t+3)		-0.0226***
		(0.0084)
R^2	0.92949	0.92938
Observations	12,150	10,827
Panel B: In-person Turnout	**(1)**	**(2)**
Treatement (t-4)	0.0133*	
	(0.0073)	
Treatement (t-3)	-0.0066	
	(0.0050)	
Treatement (t-2)	-0.0072	
	(0.0074)	
Treatement (t+0/Federal)	-0.0241***	
	(0.0091)	
Treatement (t+0/State)	-0.0222**	
	(0.0090)	
Treatement (t+1)	-0.0278***	
	(0.0074)	

Table A8: Heterogeneity by Migrant Backgrounds

	Total Turnout		
	(1)	(2)	(3)
Treatement (t-4)	0.0061	0.0059	0.0051
	(0.0062)	(0.0062)	(0.0056)
Treatement (t-3)	-0.0068	-0.0068	-0.0061
	(0.0049)	(0.0048)	(0.0044)
Treatement (t-2)	-0.0088	-0.0072	-0.0055
	(0.0055)	(0.0060)	(0.0044)
Treatement (t+0/Federal)	-0.0214***	-0.0239***	-0.0228***
	(0.0064)	(0.0066)	(0.0063)
Treatement (t+0/State)	-0.0216***	-0.0242***	-0.0230***
	(0.0065)	(0.0067)	(0.0064)
Treatement (t+1)	-0.0107*	-0.0143**	-0.0155***
	(0.0055)	(0.0060)	(0.0059)
Treatement (t+2)	-0.0160***	-0.0185***	-0.0183***
	(0.0055)	(0.0056)	(0.0056)
Treatement (t+3)	-0.0133*	-0.0139*	-0.0114*
	(0.0074)	(0.0078)	(0.0062)
Treatement (t-4) × Migrant Background	0.0024		
	(0.0070)		
Treatement (t-3) × Migrant Background	-0.0033		
	(0.0058)		
Treatement (t-2) × Migrant Background	-0.0004		
	(0.0057)		
Treatement (t+0/Federal) × Migrant Background	-0.0133*		
	(0.0077)		
Treatement (t+0/State) × Migrant Background	-0.0137*		
	(0.0078)		
Treatement (t+1) × Migrant Background	-0.0114*		
	(0.0063)		
Treatement (t+2) × Migrant Background	-0.0074		
	(0.0057)		
Treatement (t+3) × Migrant Background	-0.0110		
	(0.0084)		
Treatement (t-4) × Female		-0.0041	
		(0.0054)	
Treatement (t-3) × Female		-0.0013	
		(0.0046)	
Treatement (t-2) × Female		-0.0025	
		(0.0049)	
Treatement (t+0/Federal) × Female		0.0054	
		(0.0061)	
Treatement (t+0/State) × Female		0.0030	
		(0.0069)	
Treatement (t+1) × Female		0.0112*	
		(0.0064)	
Treatement (t+2) × Female		0.0014	
		(0.0048)	
Treatement (t+3) × Female		-0.0025	
		(0.0068)	
Treatement (t-4) × Welfare Recipient			-0.0073
			(0.0067)
Treatement (t-3) × Welfare Recipient			0.0016
			(0.0059)
Treatement (t-2) × Welfare Recipient			0.0033
			(0.0053)
Treatement (t+0/Federal) × Welfare Recipient			-0.0085
			(0.0076)
Treatement (t+0/State) × Welfare Recipient			-0.0074
			(0.0079)
Treatement (t+1) × Welfare Recipient			-0.0006
			(0.0073)
Treatement (t+2) × Welfare Recipient			-0.0108*
			(0.0057)
Treatement (t+3) × Welfare recipient			-0.0035
			(0.0077)
R^2	0.93633	0.93023	0.94409
Observations	12,150	12,150	12,150

Table A9: Heterogeneity by Age

	Total Turnout							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatement (t-4) × Age 0-6	1.96×10^{-5}							
	(0.0057)							
Treatement (t-3) × Age 0-6	-0.0109**							
	(0.0046)							
Treatement (t-2) × Age 0-6	-0.0091							
	(0.0057)							
Treatement (t+0/Federal) × Age 0-6	-0.0132**							
	(0.0060)							
Treatement (t+0/State) × Age 0-6	-0.0112*							
	(0.0061)							
Treatement (t+1) × Age 0-6	-0.0053							
	(0.0063)							
Treatement (t+2) × Age 0-6	-0.0067							
	(0.0052)							
Treatement (t+3) × Age -	-0.0170**							
	(0.0080)							
Treatement (t-4) × Age 6-18		0.0053						
		(0.0058)						
Treatement (t-3) × Age 6-18		-0.0084*						
		(0.0049)						
Treatement (t-2) × Age 6-18		-0.0085						
		(0.0064)						
Treatement (t+0/Federal) × Age 6-18		-0.0166**						
		(0.0067)						
Treatement (t+0/State) × Age 6-18		-0.0163**						
		(0.0067)						
Treatement (t+1) × Age 6-18		-0.0128**						
		(0.0062)						
Treatement (t+2) × Age 6-18		-0.0049						
		(0.0050)						
Treatement (t+3) × Age 6-18		-0.0088						
		(0.0081)						
Treatement (t-4) × Age 18-25			0.0044					
			(0.0067)					
Treatement (t-3) × Age 18-25			-0.0033					
			(0.0050)					
Treatement (t-2) × Age 18-25			-0.0019					
			(0.0047)					
Treatement (t+0/Federal) × Age 18-25			-0.0141**					
			(0.0068)					
Treatement (t+0/State) × Age 18-25			-0.0140**					
			(0.0071)					
Treatement (t+1) × Age 18-25			-0.0064					
			(0.0063)					
Treatement (t+2) × Age 18-25			-0.0121**					
			(0.0049)					
Treatement (t+3) × Age 18-25			-0.0134*					
			(0.0072)					
Treatement (t-4) × Age 25-35				0.0049				
				(0.0051)				
Treatement (t-3) × Age 25-35				-0.0075*				
				(0.0045)				
Treatement (t-2) × Age 25-35				-0.0076				
				(0.0055)				
Treatement (t+0/Federal) × Age 25-35				-0.0132**				
				(0.0056)				
Treatement (t+0/State) × Age 25-35				-0.0129**				
				(0.0056)				
Treatement (t+1) × Age 25-35				-0.0116**				
				(0.0052)				
Treatement (t+2) × Age 25-35				-0.0075				
				(0.0056)				
Treatement (t+3) × Age 25-35				-0.0171**				
				(0.0071)				
Treatement (t-4) × Age 35-45					0.0083			
					(0.0052)			
Treatement (t-3) × Age 35-45					-0.0091**			
					(0.0039)			
Treatement (t-2) × Age 35-45					-0.0116***			
					(0.0042)			
Treatement (t+0/Federal) × Age 35-45					-0.0102**			
					(0.0047)			
Treatement (t+0/State) × Age 35-45					-0.0096**			
					(0.0048)			
Treatement (t+1) × Age 35-45					-0.0075			
					(0.0053)			
Treatement (t+2) × Age 35-45					0.0012			
					(0.0049)			
Treatement (t+3) × Age 35-45					-0.0124**			
					(0.0059)			
Treatement (t-4) × Age 45-60						0.0141***		
						(0.0050)		
Treatement (t-3) × Age 45-60						-0.0156***		
						(0.0043)		
Treatement (t-2) × Age 45-60						-0.0136***		
						(0.0046)		
Treatement (t+0/Federal) × Age 45-60						-0.0084		
						(0.0055)		
Treatement (t+0/State) × Age 45-60						-0.0072		
						(0.0057)		
Treatement (t+1) × Age 45-60						-0.0181***		
						(0.0063)		
Treatement (t+2) × Age 45-60						0.0015		
						(0.0048)		
Treatement (t+3) × Age 45-60						-0.0072		
						(0.0061)		
Treatement (t-4) × Age 60-70							-0.0118**	
							(0.0051)	
Treatement (t-3) × Age 60-70							0.0143***	
							(0.0043)	
Treatement (t-2) × Age 60-70							0.0142***	

Table A10: Election Results (Vote Share)

	LINKE	The Greens	SPD	FDP	CDU	AfD
	(1)	(2)	(3)	(4)	(5)	(6)
Treatement (t-4)	0.0168	0.0296**	-0.0057	0.0463*	-0.0014	0.0244
	(0.0130)	(0.0117)	(0.0083)	(0.0261)	(0.0117)	(0.0182)
Treatement (t-3)	0.0270*	0.0200	0.0145	0.0316*	0.0107	0.0158
	(0.0152)	(0.0122)	(0.0096)	(0.0177)	(0.0126)	(0.0108)
Treatement (t-2)	0.0062	0.0148	0.0113	0.0189	0.0322**	0.0145
	(0.0135)	(0.0108)	(0.0105)	(0.0138)	(0.0140)	(0.0098)
Treatement (t+0/Federal)	0.0261	0.0339***	0.0007	0.0285**	-0.0217*	0.0074
	(0.0163)	(0.0127)	(0.0081)	(0.0143)	(0.0120)	(0.0177)
Treatement (t+0/State)	0.0490***	0.0319**	-0.0084	0.0392**	-0.0134	-0.0020
	(0.0182)	(0.0134)	(0.0093)	(0.0156)	(0.0124)	(0.0173)
Treatement (t+1)	0.0629***	0.0337*	-0.0048	0.0350*	0.0026	0.0035
	(0.0203)	(0.0182)	(0.0087)	(0.0196)	(0.0144)	(0.0192)
Treatement (t+2)	0.0593***	0.0214	0.0070	0.0168	0.0131	-0.0073
	(0.0208)	(0.0138)	(0.0104)	(0.0224)	(0.0107)	(0.0155)
Treatement (t+3)	0.0175	0.0170	0.0028	0.0314*	0.0145	0.0037
	(0.0178)	(0.0147)	(0.0104)	(0.0189)	(0.0126)	(0.0138)
R^2	0.96337	0.96861	0.92351	0.93735	0.96609	0.95758
Observations	12,150	12,150	12,150	12,150	12,150	12,150
Precinct FE	✓	✓	✓	✓	✓	✓
Election-District FE	✓	✓	✓	✓	✓	✓

Notes: The table presents event study results based on Equation 5 for electoral outcomes of the State *Governing Coalition*, comprising *Social Democrats*, the *Left Party*, and the *Green Party*, the *Opposition* (*CDU, AfD, and FDP*, and for the Right-wing populist party *AfD*. The reference election ($t-1$) is the 2019 European Election. Standard errors are clustered at the postal precinct level and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.