

UNIVERSITÀ COMMERCIALE “LUIGI BOCCONI”
PHD SCHOOL

PhD program in: Public Policy and Administration

Cycle: 35

Disciplinary Field (code): SPS/04

**Electoral Accountability for Illiberal Governments:
Essays on Immigration, State-Sponsored
Homophobia and Political Methodology**

Advisor: Catherine E. DE VRIES

PhD Thesis by
Konstantin BOGATYREV

ID number: 3106691

Year: 2025

Abstract

Far-right parties have surged to power across democracies, but their electoral accountability remains poorly understood. Through a series of essays, this dissertation provides a comprehensive analysis of how voters respond to the policies and outcomes of far-right incumbents and how policy effects on different parties can be evaluated. The findings contribute to the literature on political behavior, illiberal and far-right politics, and political methodology.

In the [first chapter](#), I empirically investigate to what extent performance in office affects support for the far right, and which voters hold them accountable. To answer these questions, I conducted a novel pre-registered survey experiment embedded in a representative survey of 2055 individuals in Italy. Respondents were randomly exposed to factual information about rising immigration or economic growth under the far-right government of Giorgia Meloni. I find that rising immigration erodes support towards the government, with an effect comparable to economic voting. The effect does not vary by partisanship, policy preferences, or issue salience, suggesting a broad-based valence reaction rather than a partisan or ideological one. These findings suggest that far-right governments are not insulated from electoral accountability, as immigration can influence voter support in a manner comparable to economic conditions.

In the [second chapter](#), I study whether state-sponsored homophobia bolsters the support for the government. While a growing body of literature documents the increasing politicization of LGBTQ- and gender-related issues by illiberal elites, little is known about the electoral effects of these strategies. I address this important question by studying whether anti-LGBTQ mobilization pays off electorally for the initiating party. Empirically, I study the adoption of anti-LGBTQ resolutions in many Polish municipalities prior to the 2019 parliamentary election. Using a synthetic difference-in-differences design, I

find that these resolutions significantly depressed turnout in affected municipalities, with opposition parties showing less mobilization capacity. By contrast, turnout for the incumbent Law and Justice Party increased substantially. Overall, this study's findings are relevant for understanding the electoral consequences of both elite-led mobilization against stigmatized and discriminated groups, and policies of subnational democratic backsliding.

In the [third chapter](#), I address the methodological challenge of estimating causal effects on party vote shares by synthetic control methods. Synthetic control methods are widely used for causal inference in case studies and panel data settings, often applied to model counterfactuals for proportional outcomes. However, conventional synthetic control methods are designed for univariate outcomes, leading researchers to model counterfactuals for each proportion separately. This study introduces an extension to synthetic control methods to simultaneously handle multivariate outcomes in compositional data. The approach establishes constant control comparisons by using the same weights for each proportion, improving comparability while adhering to treatment constraints. Results from a simulation study and an empirical application of the method to the study from chapter 2 and another study on climate policy effects underscore the benefits of accounting for the interplay of proportional outcomes. This advancement extends the validity and reliability of synthetic control estimates to common outcomes in political science.

*To Denis,
for being there during this journey.*

Acknowledgements

Many doctoral students struggle to find their way through the dissertation journey, and I was no exception. Fortunately, the support and guidance of my mentors, colleagues, friends, and family have illuminated my path, shaping not only my research but the scholar behind it. While it would be impossible to fully thank everyone who made a difference in a few pages, I would like to express my deepest gratitude to those whose support was particularly crucial to the completion of this dissertation.

It is difficult to overstate the impact my PhD advisor, Catherine de Vries, has had on my academic journey. Without Catherine, I would not have become the political scientist I am today. Coming from an economics background but deeply drawn to political science, I was incredibly fortunate to have an advisor who was ready and willing to believe in me and my ideas – and did so even in the darkest times. Catherine was always there when it mattered most, patiently teaching me the craft of academic work, serving as an excellent role model, and offering support far beyond the requirements of supervision. Through her ERC-funded project, I was able to implement the survey experiment described in the first chapter of this dissertation and to present my work at EPSA and APSA conferences. Catherine’s mentorship has been both a privilege and a blessing.

Beyond finding a wonderful PhD advisor, my academic home at Bocconi has given me the chance to work alongside and learn from brilliant scholars, whose feedback and encouragement have greatly enriched my research. Warm thanks go to Lanny Martin and Scott Williamson for their support and thoughtful comments on my PhD thesis prospectus. I have also benefited from feedback and questions in seminars and meetings with Ala Alrababah, Luca Bellodi, Maria Carreri, Simone Cremaschi, Giovanna Invernizzi, Kerim Kavakli, Dirck de Kleer, Nina Nikiforova, Piero Stanig, and many others. Finally, I thank Bocconi and Fondazione Invernizzi for their generous financial support.

Outside Bocconi, I have been fortunate to collaborate with and learn from many generous scholars. I thank my coauthors Lukas Stoetzer, Violeta Haas, Tarik Abou-Chadi and Heike Klüver, whose joint work with me is included in this dissertation and whose working styles taught me a great deal. I am also grateful to Elias Dinas and Tom Pepinsky for serving as external reviewers on my PhD committee and offering insightful comments and suggestions. Furthermore, valuable feedback was provided by Vin Arceneaux, Rafaela Dancygier, Stuart Turnbull-Dugarte, Bryn Rosenfeld, and many conference participants.

I am also grateful to the colleagues and institutions that welcomed me for academic visits and presentations throughout my dissertation research. My semester at Cornell University was made possible thanks to Tom Pepinsky, who taught me a lot in regular meetings and classes and has remained a supportive mentor since my time at Cornell. My two months at Humboldt University laid the groundwork for what became the second chapter of this dissertation and sparked the idea behind the third. The final lines of the introduction were also written in Germany during a four-week visit organized by the University of Mannheim. Presenting my research at conferences was also possible due to partial expense support from the Southern Political Science Association (SPSA), the American Political Science Association (APSA), the Society for Political Methodology (PolMeth) and the Institute for Human Studies (IHS).

Yet none of this would have been possible without the unwavering encouragement of my loved ones. My mother steadily supported my choice to pursue a PhD, even though distance kept us apart. My longtime friends Sonya, Misha, Nastya, Vanya and Kolya helped me keep perspective and reminded me of life beyond academia. Friends made during the PhD – Alireza, Alonso, Georgy, Hande, Ireri, Jorge, Michal, Nina, Paolo, and Vio – stood by me through both joyful and difficult moments, turning this chapter of my life into something I never had to face alone. Finally, I thank Denis, who was my partner and closest companion for over three years, and whose love, faith in me, and persistent care sustained me throughout that period. Without Denis, I would never have endured this dissertation journey, and to Denis I dedicate this text.

Contents

Abstract	1
Acknowledgements	4
Table of Contents	6
Introduction	9
Far-Right Rule: What's at Stake for Voters?	9
Theoretical Framework	11
Cases and Approaches	13
Contributions	15
Implications and Future Directions	17
References	24
1 When Immigration Hurts: Accountability for Far-Right Governments	25
1.1 Introduction	26
1.2 Electoral Accountability and the Far Right	29
1.3 Partisanship, Immigration Attitudes and the Far Right	33
1.4 Case: Italy under Meloni	36
1.5 Data and Methods	38
1.5.1 Sample and Data	38
1.5.2 Variables	38
1.5.3 Survey Experiment	40
1.5.4 Estimation	44
1.6 Results	45
1.6.1 Average Treatment Effects	45
1.6.2 Conditional Average Treatment Effects	48
1.7 Discussion	53
References	61

1.A	Appendix	62
1.A.1	Covariate Coding	62
1.A.2	Covariate Balance Assessment	63
1.A.3	Post-Treatment Questions	67
1.A.4	Creating an index of government support	71
1.A.5	Main results disaggregated by outcome items	72
1.A.6	Testing the equivalence of Vignettes 1 and 2	75
1.A.7	Sentiment towards the incumbent	77
1.A.8	Conditional Average Treatment Effects	79
2	Exploiting Divisions: Political Gains from State-Sponsored Homophobia	91
2.1	Introduction	92
2.2	Background	95
2.2.1	Political Homophobia in Poland	95
2.2.2	Anti-LGBTQ Resolutions	96
2.2.3	LGBTQ issues in the 2019 Elections	100
2.3	Data and Method	101
2.3.1	Data on Anti-LGBTQ Resolutions	101
2.3.2	Treatment	102
2.3.3	Outcomes	103
2.3.4	Sample	104
2.3.5	Identification Strategy	104
2.4	Results	108
2.4.1	The Electoral Implications of Anti-LGBTQ Resolutions	108
2.4.2	Additional Analysis and Robustness Checks	109
2.4.3	Individual Level Evidence	111
2.5	Conclusion	113
	References	118
2.A	Appendix	119
2.A.1	Parliamentary Election 2019: Anti-LGBTQ Campaign	119
2.A.2	Party Positions on LGBTQ Issues in 2019	121
2.A.3	Synthetic Control Weights	123
2.A.4	Restricted Sample	124
2.A.5	Full sDiD Output	125
2.A.6	Effects on Vote Shares and Turnout by Party	126
2.A.7	Varying Treatment Intensity	128

2.A.8	Alternative Treatment Definitions	132
2.A.9	Alternative Control Sample	137
2.A.10	Additional Robustness Checks	139
2.A.11	Municipality Count by Treatment Status & Sample	144
3	Reciprocal Shifts: Synthetic Control Methods for Proportions	145
3.1	Introduction	146
3.2	Synthetic Control Methods: Current Approaches	148
3.3	Compositional Effect Issues with Synthetic Control Estimates	150
3.3.1	A Motivating Example	150
3.3.2	Constant Control Comparison	152
3.3.3	Common Synthetic Control Weights Across Outcomes	155
3.4	Synthetic Control Methods for Proportions	157
3.4.1	Estimation with Constant Control Weights	157
3.4.2	Inference	160
3.5	Simulation Study	161
3.6	Empirical Applications	165
3.6.1	Application 1: Climate Policy and Voting Outcomes	165
3.6.2	Application 2: Electoral Effects of Anti-LGBTQ Policies	168
3.7	Discussion	171
	References	177
3.A	Appendix	178
3.A.1	Formalization of Synthetic Control Methods	178
3.A.2	Proofs	181
3.A.3	Algorithm for Synthetic control methods with common weights	183
3.A.4	Simulation	183

Introduction

Far-Right Rule: What's at Stake for Voters?

Illiberalism, characterized by a rejection of pluralism, erosion of democratic norms, and the embrace of authoritarian practices, has gained momentum across the globe in recent decades. This trend is evident in various countries, from Eastern Europe (Binev, 2023; Pirro & Stanley, 2022) and Latin America (Mainwaring & Pérez-Liñán, 2023) to parts of Asia (Massicard, 2023; Pepinsky, 2020; Varshney, 2022) and even within established Western democracies (De Vries and Hobolt, 2020, Ch. 1; Norris and Inglehart, 2019, Ch. 1). While scholars debate whether there is a global democratic decline (Knutsen et al., 2024; Little & Meng, 2024), at the very least, illiberal leaders and parties present a pressing challenge to pluralism and liberal democracy by engaging in nationalist and exclusionary rhetoric and policies.

Far-right parties and politicians are at the forefront of this illiberal wave. Defined by their opposition to the principles of liberalism, these actors reject individualism, pluralism, egalitarianism, humanitarianism, multiculturalism and cosmopolitanism (Mudde, 2019, see Introduction: Terminology; Pirro, 2023). Due to radical positions on the sociocultural “second” dimension and their promotion of authority and hierarchy, the far right is inherently authoritarian (Kitschelt, 2018; Mudde, 2007; Norris & Inglehart, 2019). With *nativism* at the core of the far-right ideology, it prioritizes the interests of the (constructed) native majority over those of marginalized groups and minorities. Depending on the political context, those groups “othered” by nativists as opposed to the imagined national community include not just immigrants but also other social groups based on race, ethnicity, gender identity and sexual orientation, religion or political preferences, and thus can include both minority groups within the native population and the international

community (Duyvendak et al., 2022; Golder, 2016). Exploiting divisions and anxieties toward social groups, far-right parties not only inherently oppose equality and civil liberties but also advocate for discriminatory, exclusionary and punitive policies (Pirro, 2023).

In light of this, understanding the consequences of the far-right in government is crucial for modern democracies, where illiberal actors increasingly attain or come close to attaining political power. Recent prominent examples include heads of national governments in Brazil, Hungary, India, Israel, Italy, Poland and the United States. Together with participation in national governing coalitions and subnational governance in a larger number of countries (Akkerman et al., 2016; Paxton, 2023), this signifies the entrance of the far right into the political mainstream, where they can decide on policies and implement their political vision. However, the lion's share of research treats the far right as a dependent variable, focusing on explaining their resurgence rather than understanding its consequences (Mayer, 2020).

How do voters respond to far-right governance? This pivotal question lies at the heart of this dissertation. With its distinctive nativist and authoritarian ideologies and exclusionary political projects, a far-right government can theoretically have substantial implications for political behavior via rhetoric and policies that affect people's lives. Far-right leaders can slow down economic growth (Funke et al., 2023), decrease the quality of bureaucracy and allow for deregulation (Bellodi et al., 2024; Röth et al., 2018) and contribute to retrenching welfare such as unemployment benefits, social assistance or family benefits (Afonso & Papadopoulos, 2015; Chueri, 2021). They also pursue illiberal reforms that weaken the independence of courts, the media and minority rights (Mudde, 2019; Pirro & Stanley, 2022). However, the consequences of the far right's policy shifts for political behavior are not yet fully understood.

This dissertation aims to shed light on this question by examining how far-right governments are held accountable by voters, particularly in the context of immigration and LGBTQ rights. Both issues are weaponized by the modern far right for their exclusionary agenda in Western societies. They argue that immigration threatens national identity

and social cohesion, thereby gaining electoral support by promising stricter immigration controls and protectionist policies (Golder, 2016; Ivarsflaten, 2008; Kitschelt, 2018). In Chapter 1, I investigate whether immigration continues to benefit far-right parties once they have assumed power.

In addition to immigrants, in many cases far-right parties also target LGBTQ communities, using homophobic and transphobic rhetoric to galvanize support and justify discriminatory policies (Ayoub & Stoeckl, 2024; Graff & Korolczuk, 2023). In Chapter 2, I study the electoral effects of state-sponsored homophobia, showing that it can benefit the far-right government while the opposition loses. Finally, in Chapter 3, I address the methodological issue of correctly estimating the effects on different parties via synthetic control methods.

Theoretical Framework

Far-right parties have transitioned from opposition to government in multiple democracies, yet theories of electoral accountability have largely focused on mainstream parties. This dissertation examines how far-right incumbents are held accountable, bridging the theories of retrospective voting and issue voting to analyze voter behavior and public opinion.

Understanding the electoral accountability for illiberal far-right governments requires examining how voters respond to issues and policies spearheaded by the far right. Issue voting theory posits that voters make decisions based on specific policy issues that are particularly salient to them (Weldon & McNeney, 2019). Salient issues can mobilize voter turnout (Kinder & Kam, 2010; Matthews, 2019) and also can activate social and political group divisions, with wedge issues persuading partisans to vote against their usual party alignment (Hillygus & Shields, 2008; Kinder & Kam, 2010). For instance, immigration and identity politics transcend the conventional left-right economic cleavages, facilitating issue-based voting and reshaping party competition and voter alignments (De Vries & Hobolt, 2020; Hellwig, 2014; Ivarsflaten, 2008; Kriesi et al., 2012; van de Wardt et al., 2014). Far-right parties have capitalized on these divisions to build electoral support,

but it remains unclear whether they face electoral consequences when in power.

In the context of support for the far right, the literature has so far focused on prospective voting, analyzing how their political platforms shape electoral support when they act as challenger parties without government experience (De Vries & Hobolt, 2020). For instance, far-right parties often highlight sociocultural issues like national identity and immigration during campaigns, making these issues central to their appeal (Golder, 2016; Ivarsflaten, 2008; Kitschelt, 2018). This can mobilize voter turnout and influence electoral choices, demonstrating the power of issue salience in shaping political preferences (Kinder & Kam, 2010; Matthews, 2019). However, when the far right transitions to governance, the perspective shifts to retrospective voting.

Shifting the focus to the far right in government, the theoretical framework of retrospective accountability becomes essential. Research consistently shows that individuals rely on retrospective voting, judging the incumbents by their records in office (Achen & Bartels, 2016; Ashworth, 2012; Healy & Malhotra, 2013; Lewis-Beck & Stegmaier, 2019). Research on retrospective voting has overwhelmingly focused on economic performance, amassing evidence consistent with both *pocketbook voting* based on individual-level outcomes (Healy et al., 2017; Steinberg, 2022; Visconti, 2019) and *sociotropic voting* based on aggregate, macro-level conditions (de Benedictis-Kessner & Warshaw, 2020; Lewis-Beck & Stegmaier, 2019; Nadeau et al., 2013). However, less is known about whether far-right incumbents are held accountable in a similar manner as mainstream parties, and based on similar issues.

Electoral accountability for far-right governments may extend beyond traditional economic evaluations, as their governance is often defined by identity-based policies and exclusionary agendas. Indeed, the issue voting theory complements retrospective accountability by emphasizing that voters make decisions based on specific policy domains that are particularly salient to them (Weldon & McNeney, 2019). While the importance of economic voting is well-established, it does not preclude the relevance of other performance areas in voter evaluations, with research showing that voters factor in multiple consid-

erations beyond economic outcomes. Voters also assess incumbents based on a range of other crucial factors, such as military casualties (Kriner & Shen, 2010), corruption levels (Klašnja et al., 2016), and the quality of public service provision (Adiguzel et al., 2023). Although more politically sophisticated voters tend to be more responsive to government performance across various domains, this gap in responsiveness diminishes when the issues at stake are particularly salient to the electorate (De Vries & Giger, 2014). This broadens the scope of retrospective evaluations beyond purely economic measures to include a variety of policy outcomes that significantly impact voter behavior, including far-right policies.

In government, far-right parties have the ability to shape public discourse and implement policies that reflect their ideological agenda. These policies include restrictive immigration laws, anti-LGBTQ legislation, and nationalist cultural reforms (Ayoub & Stoeckl, 2024; Graff & Korolczuk, 2023; Lutz, 2019). Even absent particular nativist and exclusionary policies, the mere presence of the far right in government invites heightened scrutiny and electoral responses on these issues. This dual framework of retrospective accountability and issue voting provides a structured approach for understanding the electoral consequences of far-right rule.

Cases and Approaches

This dissertation examines electoral accountability under far-right governments through the cases of Italy and Poland. As the two largest European democracies governed by far-right parties in their time, both represent high-profile instances of far-right parties achieving sustained executive power. While both cases reflect the broader rise of far-right politics in Europe, they also represent distinct political contexts, allowing for comparative insights into the mechanisms of voter accountability.

Unlike contexts where far-right parties are junior coalition partners, these cases allow us to study their rule when they directly control policymaking. This minimizes confounding effects from centrist or moderate coalition partners, ensuring that voter re-

sponses reflect assessments of far-right governance rather than broader coalition dynamics. In Italy, the far-right Brothers of Italy (*Fratelli d'Italia*, FdI) lead a majority coalition with other right-wing parties and dominate its policy agenda. Similarly, in Poland, Law and Justice (*Prawo i Sprawiedliwość*, PiS) controlled both the executive and legislative branches, as well as a large number of local governments. This contrasts with cases where coalition dynamics weakens responsibility attribution.

While Italy and Poland share common advantages as cases, they also provide meaningful variation that strengthens the dissertation's comparative perspective. While Italy is a long-established Western parliamentary democracy, Poland is a younger Central and Eastern European semi-presidential democracy that experienced democratic backsliding under Law and Justice. Moreover, Italy's case focuses on immigration, one of the most politically salient issues for far-right parties across Western industrialized countries. Poland's case, in turn, centers on state-sponsored homophobia, a relevant dimension of illiberal governance all over the world. Thus, the selected cases offer empirical leverage over different political domains.

Moreover, the dissertation captures voter attitudes in Italy 1.5 years into Meloni's tenure, providing insights into early-stage accountability. In Poland, the study examines PiS's reelection attempt, allowing an evaluation of far-right electoral performance after prolonged governance.

To study these cases, the dissertation employs three distinct methodological approaches, each suited to the research question at hand:

Chapter 1 employs a survey experiment embedded in a nationally representative survey of Italian voters to assess whether factual information on immigration and economic growth shifts support for the far-right government. This experimental design isolates causal effects by controlling for confounding political dynamics.

Chapter 2 uses a synthetic difference-in-differences approach to estimate the electoral impact of state-sponsored homophobia in Poland. By constructing counterfactuals from untreated municipalities, this method provides a rigorous test of how discriminatory

policies influence voter support for PiS and opposition parties.

Chapter 3 advances synthetic control methods by addressing a key limitation in estimating multivariate proportional outcomes, such as party vote shares. This refinement strengthens causal inference in electoral studies and is applied both to the Poland study from Chapter 2 and a case on the 2019 Spanish general election.

Together, these cases and methodological approaches provide a rigorous framework for analyzing how voters respond to far-right governance across different issue domains and institutional contexts, combining experimental precision with quasi-experimental causal inference.

Contributions

The first chapter of this dissertation examines whether and how far-right governments face electoral accountability for their policy outcomes, particularly in domains central to their political agenda. While extensive research has explored the conditions under which rising immigration increases support for far-right parties in opposition (Cools et al., 2021; Gessler & Hunger, 2022; Tolsma et al., 2021), less is known about how voters hold them accountable for immigration policy outcomes. To address this question, I conducted a novel pre-registered survey experiment embedded in a nationally representative survey of Italian voters (N=2055), where respondents were randomly exposed to factual information about immigration or economic growth under the far-right government of Giorgia Meloni.

The findings demonstrate that far-right governments are not insulated from electoral accountability: rising immigration significantly reduces government approval, perceived competence, and vote intention for the governing parties. Notably, this effect is comparable in magnitude to economic voting, suggesting that voters highly salient policy areas can be no less important than economic performance. Moreover, the effect does not vary by partisanship, policy preferences, or issue salience, indicating a broad-based valence reaction rather than a partisan or ideological response. This chapter contributes to research on political accountability by extending retrospective voting theories beyond

economic conditions, showing that far-right incumbents can be evaluated based on their ability to deliver on their core issues.

The [second chapter](#) examines whether state-sponsored homophobia results in electoral gains for far-right governments. Specifically, it investigates how a nationwide campaign of local anti-LGBTQ resolutions in Poland affected the 2019 parliamentary election results. Using a synthetic difference-in-differences design, the research shows that these resolutions benefitted for the incumbent far-right Law and Justice party (PiS) while reducing the overall turnout in the affected areas. Crucially, the effect was driven by opposition demobilization in affected areas, while the far-right incumbent retained or even increased its support.

This indicates that state-sponsored homophobia can be an effective strategy for gaining electoral advantage for the far right. In addition to the literature on LGBTQ politics, this study contributes to our understanding of subnational illiberalism and authoritarianism (Gibson, [2005](#); Grumbach, [2023](#); O'Dwyer & Stenberg, [2022](#)) by showing how exclusionary policies, even when symbolic, can be politically consequential. It highlights how state-sponsored homophobia spearheaded by the far right influences voter mobilization and party competition, demonstrating that illiberal incumbents can use discriminatory policies not only to signal ideological commitment and rally their base but also to demobilize political opponents.

The [third chapter](#) advances the methodological toolkit for studying electoral accountability by addressing a key limitation in synthetic control methods: their inability to properly handle multivariate proportional outcomes, such as vote shares for multiple parties. Traditional synthetic control approaches generate separate counterfactual weights for each proportion, potentially distorting estimates and limiting interpretability. To overcome this, the chapter introduces an improved method that ensures consistency across proportional outcomes, enhancing the reliability of causal inference. Applying this approach to the findings from Chapter [2](#), the study reaffirms that state-sponsored homophobia in Poland benefitted the far-right Law and Justice party (PiS). Additionally,

the method is validated through a simulation study and applied to an empirical case of the 2019 Spanish general election, demonstrating its robustness across different electoral contexts.

Furthermore, the methodological contribution from Chapter 3 enhances the accuracy and reliability of causal inference in studies involving multivariate proportional data, making it broadly applicable to various research areas. By extending the synthetic control methodology to handle compositional data, this dissertation equips scholars with improved analytical tools to study multiparty electoral outcomes. The approach holds promise for causal inference with compositional panel, which extends beyond electoral research to studying outcomes such as budget and coalition portfolio allocation (Katz & King, 1999).

Overall, this dissertation contributes to research on political behavior, far-right politics, democratic accountability and political methodology. First, it advances the study of political behavior by providing empirical insights into how voters react to policy outcomes of far-right parties once they are in power. Second, it enriches the literature on far-right politics by demonstrating that voters evaluate far-right incumbents not only based on economic performance but also on their broader policy agendas, including exclusionary and illiberal strategies. It complements our understanding of democratic accountability under illiberal rule by illustrating that far-right parties can be both rewarded and punished for their performance. Finally, the dissertation's methodological innovation in synthetic control methods further enhances our ability to rigorously evaluate the effects of government agendas on proportional outcomes such as electoral results, supporting the substantial conclusions from the empirical analysis. Together, these contributions offer new theoretical, empirical, and methodological insights into the electoral accountability for far-right governance.

Implications and Future Directions

This dissertation contributes to our understanding of electoral accountability under far-right governments and highlights the role of discriminatory policies in shaping political

behavior in contemporary democracies. Its findings hold relevance not only for academic scholarship but also for policymakers and civil society actors concerned with safeguarding democracy and pluralism.

Far-right parties frequently deflect responsibility for policy failures by blaming external actors – immigrants, elites, international institutions – for policy failures, potentially shielding themselves from electoral punishment. However, this dissertation challenges the notion that far-right governments are immune to accountability. Instead, it demonstrates that even their own supporters react negatively to unmet policy expectations, particularly in domains central to the far right, such as immigration. The very issues that propel far-right parties to power may also become the ones where they face electoral consequences.

In a troubling manner, the results also suggest that voters can indeed reward exclusionary policies, as evidenced by the case study of state-sponsored homophobia in Poland in Chapter 2. This indicates that far-right parties can effectively mobilize their voter base and potentially sway undecided voters through divisive and discriminatory strategies (Abou-Chadi et al., 2021; Corrales, 2019; Payne & Santos, 2020). Such findings raise important concerns for liberal democracy, as they suggest that exclusionary strategies can reinforce illiberal governance rather than trigger electoral backlash. Furthermore, Chapter 1 demonstrates that far-right governments may also be punished for being perceived as insufficiently restrictive on issues like immigration, underscoring the self-reinforcing nature of their illiberal agendas.

Taken together, these findings underscore the complex interplay between far-right governance and voter responses, pointing to mechanisms of democratic resilience in the face of rising illiberalism. As illustrated by the case of Italy, far-right governments may be electorally penalized when perceived as failing to deliver sufficiently restrictive or exclusionary policies. This paradox points to an important implication: although far-right rhetoric may initially attract voter support, maintaining this support while in power is contingent upon successfully translating exclusionary rhetoric into tangible policy outcomes.

The patterns observed in this dissertation highlight the importance of institutional safeguards, checks and balances, political opposition and civil society activism in counter-acting illiberal governance. By constraining the implementation of exclusionary policies, strong democratic institutions and organized societal resistance can reduce the electoral gains for illiberals pursuing such strategies. As a result, preserving democratic institutions and norms (Valentim, 2024) becomes central to resisting the spread of illiberal politics. Furthermore, as exclusionary far-right policies do not automatically trigger backlash, liberal opposition-led reframing of key issues may become crucial. This dissertation points toward a critical need for further research examining which institutional and societal conditions most effectively limit the electoral appeal of exclusionary governance strategies.

These findings are most likely to generalize to cases where elite strategies, institutional constraints, and electoral dynamics resemble those observed in the examined cases. The cases of Giorgia Meloni's government in Italy and Law and Justice's rule in Poland, while highly relevant due to their political prominence, also reflect institutional characteristics that shape electoral accountability. Both multiparty systems allow voters to shift support across multiple alternatives – an accountability dynamic that differs from two-party systems. Neither country has a presidential system, meaning executive power is more constrained than in contexts where far-right leaders wield stronger authority.

Beyond institutional factors, far-right parties and politicians vary significantly across contexts in both their ideological emphasis and strategies for mobilizing support. Even between the examined cases, Meloni's Brothers of Italy prioritized an anti-immigration platform, while Law and Justice focused more on an anti-gender and anti-LGBTQ agenda. These differences illustrate how far-right incumbents may tailor their exclusionary rhetoric and policy priorities to domestic political landscapes. As the largest European democracies governed by far-right parties in their time, Italy and Poland provide critical test cases, but further research is needed to assess whether similar dynamics hold in smaller or younger democracies, presidential systems, or non-European contexts where far-right actors may operate under different constraints and incentives.

A second limitation relates to the specific methodological approaches employed in this dissertation. The survey experiments used in Chapter 1 provide strong internal validity and also approximate real-world conditions by using factual information similar to what voters encounter in the media. Still, they may not fully replicate real-world conditions where voters are exposed to multiple and often competing sources of information simultaneously. The effects identified in a controlled environment may thus differ in magnitude when immigration information competes with other prominent political narratives or real-world partisan cues. Similarly, while causal panel data methods used in Chapter 2 provide rigorous counterfactual estimation, they depend on identifying assumptions, data availability, and stable comparability between treated and control units. Chapter 3 explicitly addresses one such methodological limitation – handling compositional data – but future research must also contend with limitations related to dynamic treatment effects, spatial spillovers, and complex interactions of multiple policy interventions over time.

In response to these limitations, future research should leverage richer datasets, longitudinal analyses, and comparative experimental designs. Extending the investigation to broader sets of cases, including comparative cross-national studies, would allow researchers to identify whether voter accountability mechanisms documented here hold universally or differ substantially under varying institutional constraints, cultural environments, or historical legacies of illiberalism. In addition, future work could benefit from combining experimental methods with observational data, thus validating the external applicability of experimentally derived insights within real-world political contexts.

Finally, subsequent work could extend investigations of electoral accountability under far-right governance to additional salient policy domains, such as gender equality, environmental policy, educational reforms, or welfare. For this purpose, future methodological extensions in causal inference could refine our understanding of the electoral consequences of illiberal governance. Furthermore, interdisciplinary work drawing from sociology, economics, and psychology could deepen insights into the cognitive and social mechanisms shaping voter behavior under illiberal rule.

References

- About-Chadi, T., Breyer, M., & Gessler, T. (2021). The (re)politicisation of gender in Western Europe. *European Journal of Politics and Gender*, 4(2), 311–314. <https://doi.org/10.1332/251510821X16177312096679>
- Achen, C. H., & Bartels, L. M. (2016). *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton University Press.
- Adiguzel, F. S., Cansunar, A., & Corekcioglu, G. (2023). Out of Sight, Out of Mind? Electoral Responses to the Proximity of Health Care. *The Journal of Politics*, 85(2), 667–683. <https://doi.org/10.1086/722040>
- Afonso, A., & Papadopoulos, Y. (2015). How the Populist Radical Right Transformed Swiss Welfare Politics: From Compromises to Polarization. *Swiss Political Science Review*, 21(4), 617–635. <https://doi.org/10.1111/spsr.12182>
- Akkerman, T., Lange, S., & Rooduijn, M. (Eds.). (2016). *Radical Right-Wing Populist Parties in Western Europe: Into the Mainstream?* Routledge. <https://doi.org/10.4324/9781315687988>
- Ashworth, S. (2012). Electoral Accountability: Recent Theoretical and Empirical Work. *Annual Review of Political Science*, 15(1), 183–201. <https://doi.org/10.1146/annurev-polisci-031710-103823>
- Ayoub, P. M., & Stoeckl, K. (2024). The Global Resistance to LGBTIQ Rights. *Journal of Democracy*, 35(1), 59–73. <https://doi.org/10.1353/jod.2024.a915349>
- Bellodi, L., Morelli, M., & Vannoni, M. (2024). A Costly Commitment: Populism, Economic Performance, and the Quality of Bureaucracy. *American Journal of Political Science*, 68(1), 193–209. <https://doi.org/10.1111/ajps.12782>
- Binev, B. S. (2023). Post-Communist Junctures, the Left, and Illiberalism: Theory with Evidence from Central and Eastern Europe. *Comparative Political Studies*, 56(4), 465–502. <https://doi.org/10.1177/00104140221109432>
- Chueri, J. (2021). Social Policy Outcomes of Government Participation by Radical Right Parties. *Party Politics*, 27(6), 1092–1104. <https://doi.org/10.1177/1354068820923496>
- Cools, S., Finseraas, H., & Rogeberg, O. (2021). Local Immigration and Support for Anti-Immigration Parties: A Meta-Analysis. *American Journal of Political Science*, 65(4), 988–1006. <https://doi.org/10.1111/ajps.12613>
- Corrales, J. (2019). The Expansion of LGBT Rights in Latin America and the Backlash. In M. J. Bosia, S. M. McEvoy, & M. Rahman (Eds.), *The Oxford Handbook of Global LGBT and Sexual Diversity Politics* (pp. 185–200). Oxford University Press.

- De Vries, C. E., & Giger, N. (2014). Holding governments accountable? Individual heterogeneity in performance voting. *European Journal of Political Research*, 53(2), 345–362. <https://doi.org/10.1111/1475-6765.12033>
- De Vries, C. E., & Hobolt, S. B. (2020). *Political Entrepreneurs: The Rise of Challenger Parties in Europe*. Princeton University Press.
- de Benedictis-Kessner, J., & Warshaw, C. (2020). Accountability for the Local Economy at All Levels of Government in United States Elections. *American Political Science Review*, 114(3), 660–676. <https://doi.org/10.1017/S0003055420000027>
- Duyvendak, J. W., Kešić, J., & Stacey, T. (2022). Nativism: What Is Returning, Where, and Why? In J. W. Duyvendak, J. Kesic, & T. Stacey (Eds.), *The Return of the Native: Can Liberalism Safeguard Us Against Nativism?* (pp. 1–26). Oxford University Press. <https://doi.org/10.1093/oso/9780197663035.003.0001>
- Funke, M., Schularick, M., & Trebesch, C. (2023). Populist Leaders and the Economy. *American Economic Review*, 113(12), 3249–3288. <https://doi.org/10.1257/aer.20202045>
- Gessler, T., & Hunger, S. (2022). How the Refugee Crisis and Radical Right Parties Shape Party Competition on Immigration. *Political Science Research and Methods*, 10(3), 524–544. <https://doi.org/10.1017/psrm.2021.64>
- Gibson, E. L. (2005). Boundary Control: Subnational Authoritarianism in Democratic Countries. *World Politics*, 58(1), 101–132. <https://doi.org/10.1353/wp.2006.0018>
- Golder, M. (2016). Far Right Parties in Europe. *Annual Review of Political Science*, 19(1), 477–497. <https://doi.org/10.1146/annurev-polisci-042814-012441>
- Graff, A., & Korolczuk, E. (2023). Gender and Illiberal Politics. In M. Laruelle (Ed.), *The Oxford Handbook of Illiberalism*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197639108.013.12>
- Grumbach, J. M. (2023). Laboratories of Democratic Backsliding. *American Political Science Review*, 117(3), 967–984. <https://doi.org/10.1017/S0003055422000934>
- Healy, A., & Malhotra, N. (2013). Retrospective Voting Reconsidered. *Annual Review of Political Science*, 16(1), 285–306. <https://doi.org/10.1146/annurev-polisci-032211-212920>
- Healy, A., Persson, M., & Snowberg, E. (2017). Digging into the Pocketbook: Evidence on Economic Voting from Income Registry Data Matched to a Voter Survey. *American Political Science Review*, 111(4), 771–785. <https://doi.org/10.1017/S0003055417000314>
- Hellwig, T. (2014). The Structure of Issue Voting in Postindustrial Democracies. *Sociological Quarterly*, 55(4), 596–624. <https://doi.org/10.1111/tsq.12072>
- Hillygus, D. S., & Shields, T. G. (2008). *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton University Press.
- Ivarsflaten, E. (2008). What Unites Right-Wing Populists in Western Europe?: Re-Examining Grievance Mobilization Models in Seven Successful Cases. *Comparative Political Studies*, 41(1), 3–23. <https://doi.org/10.1177/0010414006294168>

- Katz, J. N., & King, G. (1999). A Statistical Model for Multiparty Electoral Data. *The American Political Science Review*, 93(1), 15–32. <https://doi.org/10.2307/2585758>
- Kinder, D. R., & Kam, C. D. (2010). *Us Against Them: Ethnocentric Foundations of American Opinion*. University of Chicago Press.
- Kitschelt, H. (2018). Party Systems and Radical Right-Wing Parties. In J. Rydgren (Ed.), *The Oxford Handbook of the Radical Right*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190274559.013.9>
- Klašnja, M., Tucker, J. A., & Deegan-Krause, K. (2016). Pocketbook vs. Sociotropic corruption voting. *British Journal of Political Science*, 46(1), 67–94. <https://doi.org/10.1017/S0007123414000088>
- Knutsen, C. H., Marquardt, K. L., Seim, B., Coppedge, M., Edgell, A. B., Medzihorsky, J., Pemstein, D., Teorell, J., Gerring, J., & Lindberg, S. I. (2024). Conceptual and Measurement Issues in Assessing Democratic Backsliding. *PS: Political Science & Politics*, 57(2), 162–177. <https://doi.org/10.1017/S104909652300077X>
- Kriesi, H., Grande, E., Dolezal, M., Helbling, M., Höglinger, D., Hutter, S., & Wüest, B. (2012). *Political Conflict in Western Europe*. Cambridge University Press.
- Kriner, D. L., & Shen, F. X. (2010). *The Casualty Gap: The Causes and Consequences of American Wartime Inequalities*. Oxford University Press.
- Lewis-Beck, M. S., & Stegmaier, M. (2019). Economic Voting. In R. D. Congleton, B. Grofman, & S. Voigt (Eds.), *The Oxford Handbook of Public Choice*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190469733.013.12>
- Little, A. T., & Meng, A. (2024). What We Do and Do Not Know about Democratic Backsliding. *PS: Political Science & Politics*, 57(2), 224–229. <https://doi.org/10.1017/S1049096523001038>
- Lutz, P. (2019). Variation in Policy Success: Radical Right Populism and Migration Policy. *West European Politics*, 42(3), 517–544. <https://doi.org/10.1080/01402382.2018.1504509>
- Mainwaring, S., & Pérez-Liñán, A. (2023). Why Latin America’s Democracies Are Stuck. *Journal of Democracy*, 34(1), 156–170. <https://doi.org/10.1353/jod.2023.0010>
- Massicard, E. (2023). Illiberalism as a key to de-democratization: The case of Turkey. In M. Laruelle (Ed.), *The Oxford Handbook of Illiberalism*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197639108.013.25>
- Matthews, J. S. (2019). Issue Priming Revisited: Susceptible Voters and Detectable Effects. *British Journal of Political Science*, 49(2), 513–531. <https://doi.org/10.1017/S0007123416000715>
- Mayer, N. (2020). Political Science Approaches to the Far Right. In *Researching the Far Right* (pp. 17–31). Routledge.
- Mudde, C. (2007). *Populist Radical Right Parties in Europe*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511492037>
- Mudde, C. (2019). *The Far Right Today*. Polity.

- Nadeau, R., Lewis-Beck, M. S., & Bélanger, É. (2013). Economics and Elections Revisited. *Comparative Political Studies*, 46(5), 551–573. <https://doi.org/10.1177/0010414012463877>
- Norris, P., & Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press. <https://doi.org/10.1017/9781108595841>
- O'Dwyer, C., & Stenberg, M. (2022). Local-Level Democratic Backsliding? The Consolidation of Aspiring Dominant-Party Regimes in Hungary and Poland. *Government and Opposition*, 57(3), 508–531. <https://doi.org/10.1017/gov.2021.12>
- Paxton, F. (2023). *Restrained Radicals: Populist Radical Right Parties in Local Government*. Cambridge University Press. <https://doi.org/10.1017/9781009379069>
- Payne, L. A., & Santos, A. A. d. S. (2020). The Right-Wing Backlash in Brazil and Beyond. *Politics & Gender*, 16(1), 32–38. <https://doi.org/10.1017/S1743923X20000057>
- Pepinsky, T. (2020). Migrants, Minorities, and Populism in Southeast Asia. *Pacific Affairs*, 93(3), 593–610. <https://doi.org/10.5509/2020933593>
- Pirro, A. L. P. (2023). Far right: The significance of an umbrella concept. *Nations and Nationalism*, 29(1), 101–112. <https://doi.org/10.1111/nana.12860>
- Pirro, A. L. P., & Stanley, B. (2022). Forging, Bending, and Breaking: Enacting the “Illiberal Playbook” in Hungary and Poland. *Perspectives On Politics*, 20(1), 86–101. <https://doi.org/10.1017/S1537592721001924>
- Röth, L., Afonso, A., & Spies, D. C. (2018). The Impact of Populist Radical Right Parties on Socio-Economic Policies. *European Political Science Review*, 10(3), 325–350. <https://doi.org/10.1017/S1755773917000133>
- Steinberg, D. A. (2022). How Voters Respond to Currency Crises: Evidence from Turkey. *Comparative Political Studies*, 55(8), 1332–1365. <https://doi.org/10.1177/00104140211060268>
- Tolsma, J., Laméris, J., & Savelkoul, M. (2021). Exposure to asylum seekers and changing support for the radical right: A natural experiment in the Netherlands. *PLOS One*, 16(2), e0245644. <https://doi.org/10.1371/journal.pone.0245644>
- Valentim, V. (2024). *The Normalization of the Radical Right: A Norms Theory of Political Supply and Demand*. Oxford University Press.
- van de Wardt, M., De Vries, C. E., & Hobolt, S. B. (2014). Exploiting the Cracks: Wedge Issues in Multiparty Competition. *The Journal of Politics*, 76(4), 986–999. <https://doi.org/10.1017/S0022381614000565>
- Varshney, A. (2022). How India's Ruling Party Erodes Democracy. *Journal of Democracy*, 33(4), 104–118. <https://doi.org/10.1353/jod.2022.0050>
- Visconti, G. (2019). Economic Perceptions and Electoral Choices: A Design-Based Approach. *Political Science Research and Methods*, 7(4), 795–813. <https://doi.org/10.1017/psrm.2017.26>
- Weldon, S., & McNeney, D. (2019). Issue Voting: Modern and Classic Accounts. In *Oxford Research Encyclopedia of Politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.776>

Chapter 1

When Immigration Hurts: Voter Accountability for Far-Right Governments¹

Far-right parties have surged to power across democracies, but their electoral accountability remains poorly understood. To what extent does performance in office affect support for the far right, and which voters hold them accountable? To answer these questions, I conducted a novel pre-registered survey experiment embedded in a representative survey of 2055 individuals in Italy. Respondents were randomly exposed to factual information about rising immigration or economic growth under the far-right government of Giorgia Meloni. I find that rising immigration erodes support towards the government, with an effect comparable to economic voting. The effect does not vary by partisanship, policy preferences, or issue salience, suggesting a broad-based valence reaction rather than a partisan or ideological one. These findings suggest that far-right governments are not insulated from electoral accountability, as immigration can influence voter support in a manner comparable to economic conditions.

¹The study received approval from the Research Ethics Committee of Bocconi University (project number RA000707) and funding from the European Research Council (ERC Grant Agreement 864687, PI: Catherine E. De Vries). I am grateful to Ala Alrababah, Vin Arcenaux, Luca Bellodi, Simone Cremaschi, Rafaela Dancygier, Catherine E. De Vries, Elias Dinas, Timm Gries, Angelo Moretti, Nina Nikiforova, Scott Williamson and the participants of seminars and workshops at Bocconi University, the European University Institute, the University of Bern, the University of Manchester, the University of Mannheim, and Vanderbilt University for their helpful suggestions.

1.1 Introduction

Since the 1990s, far-right parties have transitioned from fringe challengers to mainstream political actors, assuming power in democracies across the world.² Building on early breakthroughs in local and national representation, the far right has breached the *cordon sanitaire*, entering national government coalitions and, in some cases, securing outright majority rule. In recent decades and especially recent years, far-right politicians assumed government offices in Argentina, Brazil, India, Israel, the United States, as well as in many European countries including Austria, Croatia, Estonia, Finland, Italy, Latvia, the Netherlands, Poland, Serbia, Slovakia, Sweden and Switzerland (Golder, 2016; Paxton, 2023, p.8). Much is known about what propels the far right to power, yet once in office, whether and how they are held accountable remains far less understood.

To what extent do far-right incumbents face electoral accountability? Theories of retrospective voting suggest that voters evaluate governments based on past performance, rewarding success and punishing failure (Achen & Bartels, 2016; Healy & Malhotra, 2013; Lewis-Beck & Stegmaier, 2019). However, partisanship, issue salience, and ideology can moderate or weaken this accountability mechanism, as voters may selectively process information, rationalize failures, or prioritize partisan and ideological alignment over sanctioning incumbents (Bisgaard, 2019; Bonilla, 2024; Naurin et al., 2019; Weldon & McNeney, 2019). Yet research on accountability has historically focused on mainstream right and left incumbents in democracies, where, until recently, the far right has mostly remained an electoral challenger. Recent evidence from Hungary and Poland suggests that, although populist far-right incumbents received fewer votes when the economy performed worse, they – like non-populists – faced only minimal electoral punishment for attacks on democracy (Grzymala-Busse & Nalepa, 2022; Jacob, 2024). This raises the question of whether far-right governments, once in power, are held accountable like their mainstream counterparts – or whether their electoral base and political strategies shield them from

²Henceforth I use *far right* an umbrella concept to encompass both radical and extreme right (Mudde, 2019; Pirro, 2023). I also use *radical right* when referring in particular to the far-right governments in democratic systems such as Italy.

meaningful voter sanctions.

I study this question by focusing on a policy domain at the heart of far-right politics in Western advanced economies: immigration. Rooted in *nativism* (Golder, 2016; Kitschelt, 2018; Mudde, 2019), far-right political entrepreneurs in Europe and the US have made anti-immigration positions central to their platform (De Vries & Hobolt, 2020; Ivarsflaten, 2008; Kitschelt, 2018). Yet once in power, their restrictive rhetoric often translates into more moderate policy outcomes (Betz et al., 2019; Lutz, 2019). The far right in government may push against immigration, but they are neither necessary nor sufficient for that, enacting policies similar to those of center-right cabinets (Akkerman & de Lange, 2012; Mudde, 2013).

To test whether voters hold far-right governments accountable for immigration outcomes, I conduct a pre-registered survey experiment embedded in a representative survey of 2055 individuals in Italy. This novel experiment examines whether exposure to factual information about rising immigration erodes public support for Giorgia Meloni’s far-right government. Additionally, it evaluates whether this effect is moderated by partisanship, policy preferences and issue salience.

Italy provides a particularly appropriate setting to examine whether voters hold far-right governments accountable for immigration outcomes. In 2022, Italy became the largest Western democracy governed by the far right. Under a far-right prime minister Giorgia Meloni, a coalition of right-wing parties established a populist majority government, thus escaping the policy constraints that the far right typically face as junior coalition partners (Akkerman & de Lange, 2012; Betz et al., 2019; Lutz, 2019). Immigration was central to Meloni’s electoral campaign, where she pledged to crack down on unauthorized arrivals and tighten border controls. Given that immigration has been directly linked to the rise of the Italian far right and remains a politically salient issue (Barone et al., 2016; Cremaschi et al., 2024; lo Polito, 2023), Italy provides an ideal context to test whether voters hold the far right accountable for its core policy domain.

Supporting the logic of performance-based accountability, the findings reveal that

rising immigration can erode electoral support for the far-right government, measured via government approval, perceived competence or voting propensity. The effect is comparable in magnitude to that of economic voting, suggesting that immigration – like economic performance – serves as a key domain for holding far-right incumbents accountable. However, the effect does not significantly vary by partisanship, policy preferences, issue salience or other voter characteristics. Voters across the political spectrum respond similarly to information about rising immigration, suggesting a broad-based valence reaction rather than a partisan or issue-driven one.

This study makes three contributions to the existing literature. First, it contributes to the study of far-right politics by examining how voters respond to the far right in office. Previous research has either focused on the far right in opposition (Cremaschi et al., 2024; De Vries & Hobolt, 2020) or on the supply side of the far-right governance (Akkerman & de Lange, 2012; Albertazzi & McDonnell, 2015; Bellodi et al., 2024; Bergmann et al., 2021; Betz et al., 2019; Funke et al., 2023; Paxton, 2023). Taking one step further and shifting the focus back to voters, this study provides microfoundations for understanding the electoral accountability for far-right incumbents. By examining the demand side in the forms of voter attitudes and behavior, this research provides insights into the feedback effects of far-right performance in government. I demonstrate that policy performance is important in shaping public perceptions of these parties.

Second, this study contributes to the broader and vast literature on accountability and retrospective voting (reviewed by Ashworth, 2012; De Vries and Solaz, 2017; Healy and Malhotra, 2013; Lewis-Beck and Stegmaier, 2019) by demonstrating how immigration outcomes can causally affect voter evaluations of the incumbent and support for the government. Previous research on democratic accountability mostly focused on economic performance (Lewis-Beck & Stegmaier, 2019), with some scholars analyzing the broader impact of salient issues (e.g., De Vries and Giger, 2014). This study makes a further step by demonstrating the causal link between immigration outcomes and voter evaluations of the government during its term, extending the principles of economic voting to the realm of immigration policy. Moreover, by testing partisanship and issue salience as moderators,

I reveal that accountability for immigration operates across the electorate rather than being limited to far-right supporters or immigration-focused voters. This broadens the scope of retrospective voting theories to encompass a wider range of salient policy areas, reflecting the multifaceted nature of voter decision-making.

Finally, bridging these literatures, I contribute to the burgeoning research on immigration and the far right by demonstrating that immigration can hurt support for the far right when the far right is in government positions. While existing literature has frequently documented how increasing immigration can boost support for far-right parties when they are in opposition (Cools et al., 2021; Gessler et al., 2022; Tolsma et al., 2021), this research reveals that rising immigration can erode support for far-right incumbents. This highlights the context-dependent nature of immigration's political effects: what helps the far right in opposition can harm them in power.

The remainder of the paper is organized as follows. The next section situates the far right within broader research and theory of electoral accountability, presenting concrete propositions about the effects of immigration on attitudes towards a far-right government. I then introduce the case selection, experimental design, data and analysis of the survey experiment. The final section discusses the implications of the results.

1.2 Electoral Accountability and the Far Right

Electoral accountability, i.e., the extent to which voters hold elected officials to account, is vital for democratic theory. Research in political science and beyond has extensively demonstrated that citizens evaluate the incumbent based on previous performance (Achen & Bartels, 2016; Healy & Malhotra, 2013; Lewis-Beck & Stegmaier, 2019). For interpreting this, political agency models offer a particularly useful and widely established theoretical framework, which views voters as principals evaluating incumbents (agents) based on performance to ensure accountability. Originating from the classical view of retrospective voting as sanctioning or rewarding incumbents, modern rational choice approaches emphasize that backward-looking evaluations serve not only to judge past actions, but also

to infer the type of politicians and expectations of their future performance (Ashworth, 2012; Duggan & Martinelli, 2017).

However, voters cannot directly observe politicians' efforts or intrinsic characteristics and must rely on observable actions and outcomes to form their evaluations. Implemented policies, the fulfillment of given promises and pledges, and realized outcomes like economic growth provide voters with information about the incumbent's competence and effort. Performance can be assessed based on personal outcomes (*egocentric voting*) or aggregate, macro-level conditions (*sociotropic voting*)³. While research on accountability is largely focused on economic outcomes, theory and evidence suggest that voters can care about an entire range of issues. For instance, voters retrospectively evaluate the incumbents based on military casualties (Kriner & Shen, 2010), corruption (De Vries & Solaz, 2017) and public service provision (Adiguzel et al., 2023).

By looking back at observable outcomes, voters not only update their projections of (egocentric or sociotropic) policy benefits, but also assess whether the government fulfilled its promises and pledges, inferring the competence and trustworthiness of political actors for future terms (Matthieß, 2024, ch. 2). Evidence suggests that pledge fulfillment is important for electoral accountability, with broken promises getting punished by voters (Bonilla, 2022; Matthieß, 2020, 2022; Naurin et al., 2019; Rose & Wessels, 2019). Notably, broken promises are punished more than kept ones are rewarded (Bonilla, 2022; Naurin et al., 2019). While voters do not interpret every statement as a pledge, they expect clear commitments to be fulfilled and politicians to be held accountable for them (Krishnarajan & Jensen, 2022). Central to this research are party manifestos, which form the core of each party's agenda and the electoral mandate it claims upon entering office (Naurin & Thomson, 2020).

In this study, I extend the electoral accountability framework to far-right parties in

³Previous research based on economic performance produced both evidence consistent with sociotropic voting (Benedictis-Kessner2020; Lewis-Beck & Stegmaier, 2019; Nadeau et al., 2013) and with pocketbook voting, i.e., egocentric economic voting (Healy et al., 2017; Steinberg, 2022; Visconti, 2019). The two considerations are not mutually exclusive and can appear significant simultaneously (Healy et al., 2017; Tertychnaya et al., 2018). Moreover, empirical evidence on the significance of economic voting does not rule out that other performance domains might matter for incumbent evaluations.

government, recognizing their distinct political mandate and ideology. Far-right parties distinguish themselves from mainstream incumbents with their nativist ideology, their challenge to liberal democratic norms, and, in most cases, their populist opposition to the establishment (Rooduijn et al., 2023). When entering governments, populists risk becoming part of the establishment they denounce, though their supporters can resolve this dissonance by maintaining trust in their government while continuing to distrust traditional institutions (Castanho Silva, 2019). Institutional representation of the far right has been shown to reshape voter attitudes, fueling polarization between their supporters and opponents (Bischof & Wagner, 2019) and normalizing these parties as political actors (Valentim, 2021). Despite efforts to ostracize the far right through a *cordon sanitaire*, public support for such exclusion remains limited and contingent on the broader political context (Turnbull-Dugarte, 2024). These dynamics have direct implications for electoral accountability: if far-right incumbents are perceived as legitimate governing parties, voters may evaluate them based on standard retrospective criteria, such as policy performance.

Furthermore, the question is not just whether the far right are held accountable by the voters but also the scope of this accountability. In contrast to other parties, the far right promise radical policy shifts, particularly on issues such as minorities, immigration, law and order, gender, and national sovereignty (Mudde, 2019). In Western industrialized countries, the far right has made immigration a central focus of their platforms (De Vries & Hobolt, 2020; Ivarsflaten, 2008; Kitschelt, 2018). Furthermore, they have become strongly associated with the issue of immigration, often becoming the political owner of that issue (Kriesi et al., 2024). Rooted in nativism, their stance is explicitly anti-immigration, typically advocating for curbing immigration inflows, stricter controls, selective admissions, or shutting down particular migration pathways. Therefore, I expect electoral accountability for far-right parties in government to be particularly pronounced with respect to their record on immigration. Given the salience of immigration in their platforms, their failure to curb immigration may undermine their credibility among voters. I propose that increasing immigration makes citizens less likely to approve of their far-right government, consider it competent and vote for it again. To operationalize this

argument in observable empirical terms for later testing, I formulate a hypothesis⁴ in terms of information acquisition:

Hypothesis 1: Exposure to information about increasing immigration erodes citizens’ attitudes towards their far-right government.

Importantly, this proposition does not imply an entirely different logic of electoral accountability for far-right parties other than the one used by political science for other parties. However, it applies this theory to a previously underexplored domain of immigration⁵. In the recent decades, immigration is highly salient among voters in Western industrialized countries (Braun & Schäfer, 2022; De Vries & Hobolt, 2020; Neundorff & Adams, 2018) – including, unsurprisingly, those where far-right parties obtained electoral success and government offices. However, empirical evidence on holding the incumbents accountable for immigration is yet limited. Using European survey data, Fumarola (2021) finds that perceived higher immigration is associated with lower levels of support for incumbents. Much less is known about the causal effect of immigration on incumbent support. Thompson (2023) finds no effect of the Mariel Boatlift on the US incumbent’s support beyond the Cuban community. Gessler et al. (2022) find that the 2015 refugee crisis benefitted the far-right Jobbik in Hungary at the expense of the incumbent Fidesz party, despite Fidesz also adopting far-right positions in response to the crisis. In a hypothetical survey experiment in the US, only co-partisans penalized representatives who broke their immigration pledges compared to those who partially or fully kept them (Bonilla, 2024).

⁴All hypotheses presented in the paper were pre-registered on OSF (**Bogatyrev2024b**). A further hypothesis on pledging from the pre-analysis plan is not included in the theory section but is discussed directly in the analysis. In addition, the pre-registration includes more details such as the questionnaire, power analysis, and ethical considerations, which are omitted in the paper.

⁵Importantly, immigration is not perceived strictly through an economic lens: immigration attitudes are also influenced by cultural concerns (Hainmueller & Hopkins, 2014). Immigration is a wedge issue that is not easily aligned with the economic left-right dimension (De Vries & Hobolt, 2020; Ivarsflaten, 2008; van de Wardt et al., 2014).

1.3 Partisanship, Immigration Attitudes and the Far Right

While this study is primarily concerned with the mere existence of electoral accountability for the far right, previous studies have shown that individuals do not always respond to government performance. To respond, voters need to learn about it, attribute the responsibility to political actors and decide act upon it regardless of their partisan, ideological or in-group considerations.

A large literature shows that partisan-motivated reasoning can weaken accountability, owing to the information partisans acquire and how they rationalize it, attributing credit or blame (Bisgaard, 2019; De Vries & Solaz, 2017; Healy & Malhotra, 2013; Weldon & McNeney, 2019). However, in hypothetical experiments with specific information on which promises were kept or broken, co-partisans were in fact the only ones who penalized broken promises in the US (Bonilla, 2024). Similarly, Naurin et al. (2019) find that while co-partisans only modestly reward kept promises but punish broken ones, non-co-partisans react even more negatively the fulfillment of a pledge they oppose. Others did not find evidence of partisan-motivated reasoning in the German context (Matthieß, 2024, ch. 6). In contexts where similar information is provided to voters regardless of their partisanship (such as a controlled experiment), one can abstract from the issue of partisan information sources. This suggests that one might expect a stronger, not weaker punishment from co-partisans of the incumbent (Bonilla, 2024; Naurin et al., 2019). Thus, to test whether partisanship moderates electoral accountability for the far right, I formulate the following empirical proposition:

Hypothesis 2: Exposure to information about increasing immigration causes a stronger negative effect on attitudes towards the far-right government among the voters of its constituent parties

An alternative hypothesis for voters' responses to immigration outcomes is that exposure to information about rising immigration does not simply trigger retrospective

evaluations of government performance, but instead reminds voters of the far right's long-standing anti-immigration positions. If this is the case, we would expect stronger negative reactions to be concentrated among opposition voters. Testing Hypothesis 2 will also allow us to assess this alternative.

The extent to which voters hold far-right governments accountable may also hinge on their attitudes toward the core issues these parties champion. Issue voting theory posits that voters make decisions based on specific policy domains that are particularly salient to them (Weldon & McNeney, 2019). Salient issues can activate social and political group divisions, with wedge issues persuading partisans to vote against their usual party alignment (Hillygus & Shields, 2008; Kinder & Kam, 2010). For instance, immigration transcends the conventional left-right economic divide, facilitating issue-based voting and transforming electoral dynamics (De Vries & Hobolt, 2020; Hellwig, 2014; Ivarsflaten, 2008; Kriesi et al., 2012; van de Wardt et al., 2014). However, the research on issue voting for the far-right has so far focused on prospective voting, based primarily on their record as challenger parties in Western democracies (De Vries & Hobolt, 2020).

The theory of issue voting complements retrospective accountability by emphasizing that voters evaluate parties based on their performance on salient issues (Weldon & McNeney, 2019). When voters care deeply about a particular issue, they are more likely to update their attitudes based on government performance in that domain (De Vries & Giger, 2014). Traditionally, economic concerns have dominated voter priorities, thus shaping the focus of accountability research on economic issues (Lewis-Beck & Nadeau, 2011; Margalit, 2019). However, political conflict is not confined to a single dimension, allowing other highly salient issues, such as crises, scandals, or immigration, to structure electoral accountability. Although more politically sophisticated voters tend to be more responsive to government performance across various domains, this gap in responsiveness diminishes for issues more salient to the electorate (De Vries & Giger, 2014).

Voters may also react based on their substantive policy preferences rather than just the salience of an issue. For example, legislators who vote for policies that diverge

from their constituents' views face electoral punishment, underscoring the role of policy congruence in retrospective evaluations (Ansolabehere & Jones, 2010). According to Naurin et al. (2019), fulfilling a policy promise tends to boost approval among supporters while decreasing it among opponents, with electoral consequences depending on the overall balance of policy supporters and opponents in the electorate. In contrast, Bonilla (2022) shows that voters penalize even violations of those promises that they disagree with. In Germany, Matthieß (2024, ch. 6) finds no moderating effect on policy preferences on accountability.

To yield observable implications of issue voting for far-right government accountability, I further focus on the issue of immigration. Recent experimental evidence points out that immigration policy positions can matter to voters (Dai & Kustov, 2024). In the case of immigration, salience is not unrelated to policy positions: in fact, those who oppose immigration tend to care more about it than those who welcome it (Kustov & Landgrave, 2023).

Drawing on the issue voting theory, I formulate two testable propositions. First, accountability should be stronger among those who consider immigration a highly salient issue, as they are more likely to react to perceived government failure. Second, anti-immigration individuals should also respond negatively to rising immigration, as their policy stance makes them particularly sensitive to changes in immigration levels.

Hypothesis 3: Exposure to information about increasing immigration causes a stronger negative effect on attitudes towards the far-right government among individuals who consider immigration an important issue

Hypothesis 4: Exposure to information about increasing immigration causes a stronger negative effect on attitudes towards the far-right government among individuals who express anti-immigration attitudes

To effectively test hypotheses 1-4, I consider three key attitudes toward the govern-

ment: approval, opinion about government competence, and propensity to vote for the government parties. Government approval captures the general satisfaction or dissatisfaction with the government's performance and serves as a broad indicator of voter sentiment. Perceived government competence reflects how voters evaluate the government's ability to effectively manage policies and address issues. Propensity to vote for the government assesses support for the incumbent government in potential elections. Examining these attitudes provides a more comprehensive view of different dimensions of voter accountability: performance evaluations, general support and electoral support. In what follows, I test each hypothesis outlined above with respect to voter attitudes.

1.4 Case: Italy under Meloni

I empirically test my hypotheses through a pre-registered survey experiment conducted in Italy in early 2024. Italy provided an ideal setting for this study, being governed by the far right while also experiencing politically salient rising immigration. These conditions make Italy a crucial case for examining whether voters hold far-right incumbents accountable for immigration outcomes.

Italy was the largest Western democracy governed by the far right at the period of our survey. Sworn in October 2022, Giorgia Meloni became the first far-right prime minister in Western Europe in the 21st century, further signifying Italy's place at the forefront of the far-right rise to power. Having a comfortable parliamentary majority, the government coalition is dominated by Meloni's *Brothers of Italy* party alongside the *League* (both far-right), with minor roles given to the smaller right-wing parties (*Forza Italia* and *Us Moderates*). The latter remained largely collaborative with the policy agenda set by their senior partners, ensuring that the far-right leadership has faced no significant internal constraints. This rules out any reasonable expectations that observed effects, or lack thereof, be attributed to other party in government rather than the far-right leadership (Royed et al., [2000](#)). Thus, the Meloni government composition and dynamics allow for testing my theoretical argument about citizens' evaluations of the government.

Besides, Italy allows for testing government accountability in the domain of immigration. Like many economically advanced countries, it faced substantial immigration in recent decades. Immigration became important to the voters and entered the political domain alongside more traditional topics such as economic development. Furthermore, just like in other Western industrialized democracies, rising immigration in Italy has been directly accompanied by and led to an increase in support for the far right (Barone et al., 2016; Io Polito, 2023). Besides, public service deprivation that fueled far-right support also amplified concerns about immigration, while the far-right skillfully blamed immigrants in their rhetoric (Cremaschi et al., 2024). As immigration remained politicized in Italy after Meloni's cabinet assumed office, the case of Italy in 2024 also allows us to test whether the relationship between immigration and attitudes to the far right has changed.

The timing of my survey experiment also provides a hard test of the theoretical argument, as Meloni's government had only been in office for less than 1.5 years before the survey, and her party was never part of the national government before that. Thus, the amount of responsibility that could be attributed to her for recent changes in fundamental indicators such as immigration or economic growth is very limited. Moreover, those outcomes naturally depend on many factors beyond the government's control. However, to what extent the Italian government was able or willing to affect immigration and economic outcomes is beyond the scope of this study. Instead, I study whether the voter's attitudes towards the government respond to those policy outcomes. The limited time since the government's entry into office works against finding significant effects in my analysis, thus presenting a hard test for my main hypothesis. Finding an effect under such circumstances would suggest that a) voters begin holding far-right governments accountable even in the early stages of their tenure, and b) this effect would likely become stronger over time, as voters increasingly associate the incumbent government actions (or inaction) with current policy outcomes.

1.5 Data and Methods

I test the hypotheses outlined above through a survey experiment using textual and graphical vignettes⁶. The experiment was part of a survey conducted in Italy in February 2024.

1.5.1 Sample and Data

The participants of the survey in February 2024 were participating in the second wave of an online panel survey, which started in October 2023. The sample was recruited by the firm SWG from a representative pool of 60,000 vetted individuals living in Italy, taking the adult Italian population as the reference population in key demographic characteristics: gender (male and female), age (six strata), place of residence (three strata: Northern Italy, Central Italy, Southern Italy) and education (three strata: high, medium, low).

The 2559 respondents of the first wave were recontacted in February 2024 for the second wave, and 2055 of them responded. In each wave, the survey took about 15 minutes and was administered through the firm’s proprietary software using the CAWI (Computer Assisted Web Interview) methodology. In the analysis, we use data from both survey waves of the panel: while the experiment took place in the second wave, the first wave also contains relevant pre-treatment covariates.

1.5.2 Variables

Before the randomized experiment, respondents answered a series of pre-treatment questions about their demographics, socioeconomic conditions, political attitudes, and voting behavior. Table 1.1 characterizes the data, based on the main descriptive statistics for the pre-treatment covariates.

These variables correspond directly to questions in the survey. For some variables, I compressed several questions into one variable to keep more degrees of freedom in the later

⁶The vignettes are described below in the section [Vignette Experiment](#).

Variable	Mean	SD	Range	N
Gender (female = 1)	0.53	0.50	0 - 1	2055
Age	51.75	16.73	18 - 98	2055
Foreign-born	0.03	0.16	0 - 1	2055
Years of schooling	12.31	3.20	5 - 20	2055
Has a spouse (yes = 1)	0.60	0.49	0 - 1	2037
Has children (yes = 1)	0.63	0.48	0 - 1	2055
Has a job (yes = 1)	0.50	0.50	0 - 1	2055
Annual income (euro)	27,785	17,274	9000 - 80000	1849
Financial gain in the last year	-0.24	0.91	-2 - 2	2055
Received financial support	0.29	0.45	0 - 1	2055
Level of financial stress	2.09	1.05	0 - 4	2055
Suffered natural hazards (yes = 1)	0.58	0.49	0 - 1	2055
Anti-feminism attitudes index	0.31	0.26	0 - 1	1956
Anti-LGBTQ attitudes index	0.36	0.27	0 - 1	1899
Voted far-right 2022 (yes = 1)	0.24	0.43	0 - 1	2055
Anti-immigration views (yes = 1)	0.55	0.50	0 - 1	2004
Immigration important (yes = 1)	0.36	0.48	0 - 1	2055

Note: Descriptions of the variables are included in the first column.

Table 1.1. Descriptive statistics for the pre-treatment covariates

analysis. Specifically, each of the variables “anti-LGBTQ attitudes” and “anti-immigrant attitudes” is an index and generated by standardizing the first principle component of several relevant questions.

Finally, three covariates used in the analysis of treatment effect heterogeneity are coded as binary variables to maximize statistical power, which is always smaller for conditional effects than for the main ones. The importance of immigration issues (*mig_imp_bin*) is coded as 1 if respondents rated the issue as highly important. The variable “anti-immigrant attitudes” (*anti_imm*) takes the value of 1 if the respondent has agreed with at least one of five anti-immigrant statements provided in the survey. Voting for far-right parties in the 2022 elections (*fr_2022*) is a binary variable where respondents who voted for Brothers of Italy, League, or Italexit are coded as 1. As Italexit is a very small extra-parliamentary party, this variable largely reflects having voted for one of the far-right parties in the government.

These variables provide a comprehensive overview of the respondents’ demographic, socioeconomic, and political characteristics, ensuring a robust analysis of the treatment effects. The full description of these variables is provided in Appendix [1.A.1](#).

1.5.3 Survey Experiment

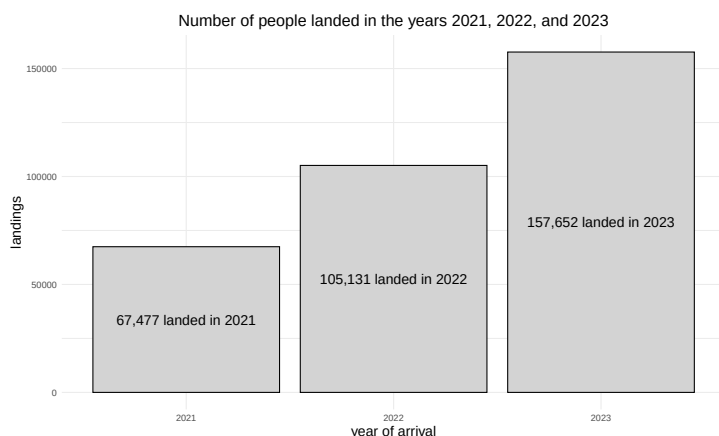
During the survey, I administered an experiment that exposed respondents in randomized groups to textual and graphical information about recent trends in immigration and economic growth. I assigned each participant to one of four groups: Treatment Group 1, Treatment Group 2, Treatment Group 3, Control Group via simple randomization. Before seeing the randomized experiment part, all respondents answered a battery of pre-treatment questions regarding their demographics, socioeconomic conditions, political attitudes and voting behavior. Participants in Group 1 are shown Vignette 1 (Immigration Outcomes), participants in Group 2 are shown Vignette 2 (Immigration Pledge and Outcomes) and participants in Group 3 are shown Vignette 3 (Economic Outcomes). Participants assigned to the Control Group directly proceed to the last part of the ques-

tionnaire without being shown any vignette.

I randomized the exposure to information about the Italian government's performance in three vignettes. Vignette 1 (Immigration Outcomes) provides Treatment Group 1 with graphical statistical information on immigrant arrivals to Italy through the Mediterranean Sea. Vignette 2 (Immigration Pledge and Outcomes) provides Treatment Group 2 with the same information together with an immigration-related statement of Giorgia Meloni from October 2022. Vignette 3 (Economic Outcomes) provides Treatment Group 3 with graphical statistical information on the Italian GDP growth. Like the entire survey, the vignettes are provided to the respondents in the Italian language. The following English version is the closest literal translation from Italian to be used in this paper and later analyses.

1. Vignette 1 (Immigration Outcomes):

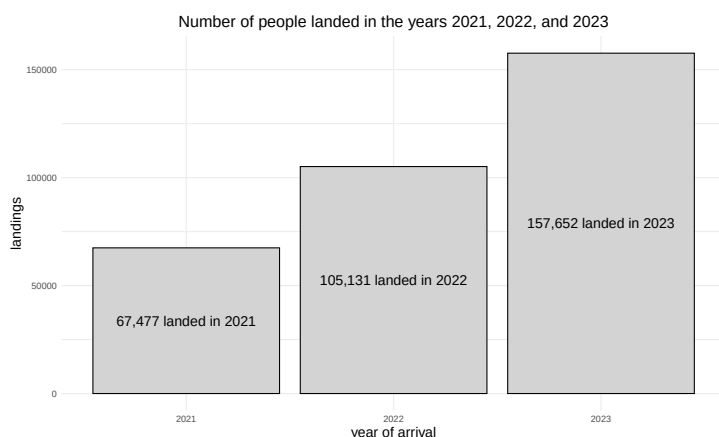
In 2023, migrant boat landings in Italy kept increasing. The graph below illustrates this, based on the government's data:



2. Vignette 2 (Immigration Pledge and Outcomes):

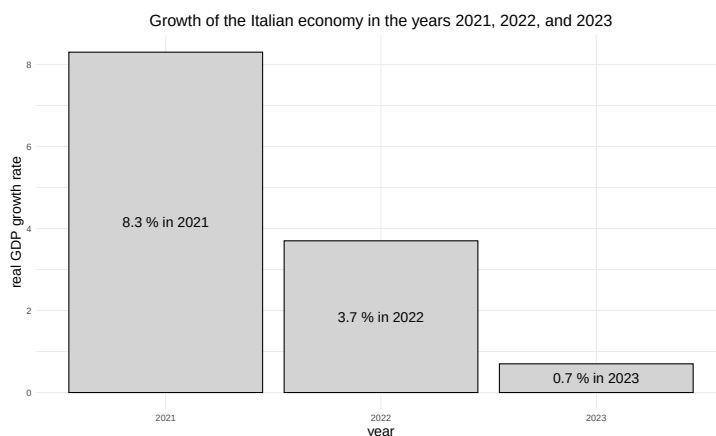
Taking office as head of the Italian government, Giorgia Meloni declared: “This government wants to follow a path, little pursued so far: stopping illegal departures, finally breaking the trafficking of human beings in the Mediterranean”

In 2023, migrant boat landings in Italy kept increasing. The graph below illustrates this, based on the government's data:



3. Vignette 3 (Economic Outcomes)⁷

In 2023, the economic growth in Italy kept slowing down. The graph below illustrates this, based on the government's data:



4. Control Group: No Treatment.

Those in the Control Group do not see any vignette. They proceed directly to the outcome questions.

I use the results from the control group to estimate the treatment effect and to compare the effects of alternative treatments above. As the control group does not receive any information, it is a pure control and not a placebo. For my experiment, the need for a placebo is very limited as the treatments provided are short and undemanding in terms of effort, mental load and physical engagement. Thus, I refrain from using a placebo in this experiment to avoid the potential effects of the placebo on the outcome (Porter & Velez, 2022). The pure control allows comparing respondents who were primed about certain policy pledges and outcomes with respondents who were not, and not with those who were distracted by some placebo information that must be completely unrelated to government performance. Thus, the pure control provides a theoretically meaningful comparison in

⁷Vignette 3 is not associated with the main hypotheses. I remain agnostic about the sign and size of the treatment effect for Vignette 3, but having it for comparison is useful for understanding the effect of Vignette 1. I.e., if Vignette 1 has no effect, will a Vignette about economy have no effect either? And if Vignette 1 has an effect, how does it compare to Vignette 3? Thus, Vignette 3 serves as a benchmark for the immigration treatment effect.

terms of potential outcomes.

1.5.4 Estimation

I use an OLS estimator and a p-value of 0.05 as a threshold of statistical significance.

I estimate average treatment effects (ATE) to explore the effect of immigration vs economic policy pledges and outcomes on attitudes. I regress each outcome variable on indicators for treatment assignment, while also controlling for covariates to improve precision.

The ATE estimates are based on the following regression:

$$Y_i = \beta_0 + \beta_1 Immig_i + \beta_2 Pled_i + \beta_3 Econ_i + \beta_4 X_i + \epsilon_i, \quad (1.1)$$

where for respondent i :

Y_i is the relevant outcome measure

$Immig_i$ equals 1 if respondent i is in treatment (vignette) group 1 or and 0 otherwise;

$Pled_i$ equals 1 if respondent i is in treatment (vignette) group 2 or and 0 otherwise;

$Econ_i$ equals 1 if respondent i is in treatment (vignette) group 3 and 0 otherwise;

$X_{i(t)}$ is a vector of relevant individual-level covariates;

$\epsilon_{i(t)}$ is a random error term;

β_0 is a constant (intercept);

β_1 is a coefficient measuring the effect of treatment (vignette) 1 on the outcome

β_2 is a coefficient measuring the effect of treatment (vignette) 2 on the outcome

β_3 is a coefficient measuring the effect of treatment (vignette) 3 on the outcome

β_4 is a vector of coefficients for the covariates included.

To aggregate information on the incumbent support, I construct an index combining

three related outcome variables: government approval, perceived government competence, and the propensity to vote for government parties. Given the conceptual overlap among these measures, I apply principal component analysis (PCA) to aggregate them into a single index corresponding to the first principle component, standardized to have a 0-1 range. The resulting government support index is very strongly ($\rho > 0.9$) correlated with each of the three outcome variables, capturing over 85% of their total variance. The PCA results are provided in Appendix 1.A.4. Subsequently, the government support index is used as the main outcome measure, while disaggregated results for the three separate outcome variables are reported in the appendix.

I introduce a battery of individual-level covariates accounting for pre-treatment socio-demographic characteristics and political metrics. The full description of these of variables is provided in Appendix 1.A.1.

1.6 Results

1.6.1 Average Treatment Effects

My theoretical framework predicts that exposure to information about increasing immigration will negatively affect citizens' attitudes toward a far-right government (hypothesis 1). The experimental results support this hypothesis for all three attitudes analyzed: government approval, perceived competence, and voting intentions. Figure 1.1 illustrates the average treatment effects (ATEs) of the experimental vignettes relative to the control group, estimated via ordinary least squares (OLS) regression. The baseline specification includes covariates listed in the Variables section because controlling for them improves precision (Clifford et al., 2021). As Figure 1.1 shows, the inclusion of covariates narrows the confidence intervals without substantially changing the point estimates of the treatment effects.

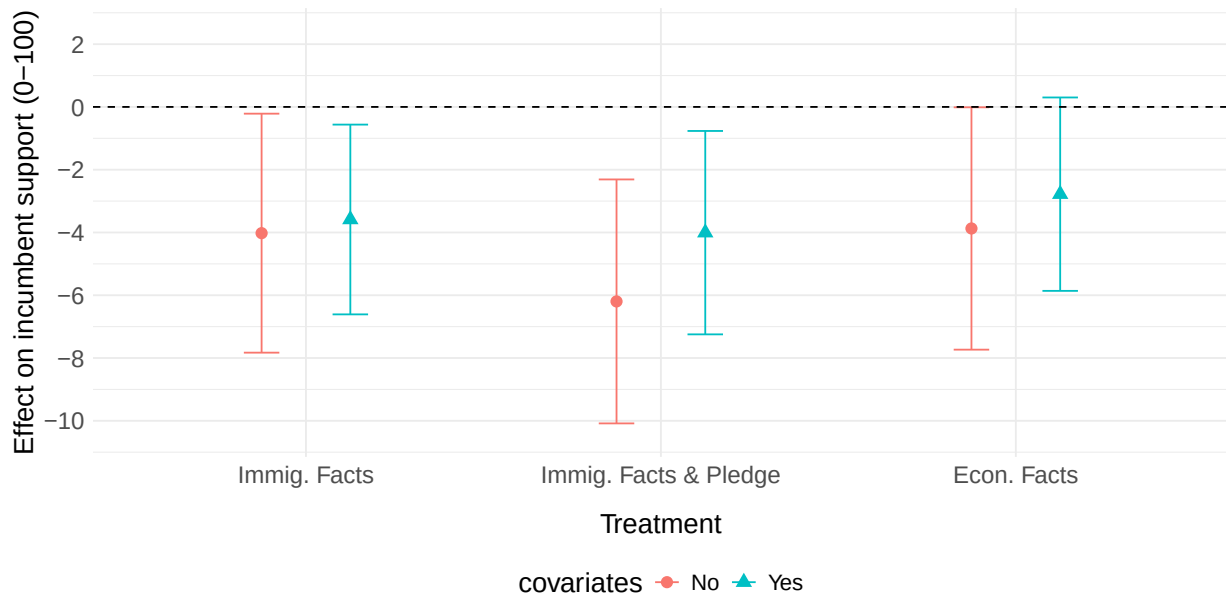


Figure 1.1. Average treatment effect of each experimental vignette relative to the control group

The results indicate that exposure to information about rising immigration decreases support for the far-right incumbent by about 4 percentage points. This effect is statistically significant at the 95% confidence level, providing strong support for **Hypothesis 1**, which posited that voters would react negatively to rising immigration under a far-right government. This suggests that the public holds the government accountable for its failure to reduce immigration levels, an issue central to its electoral mandate.

When the treatment combines immigration facts with an explicit reference to a broken pledge, the estimated effect remains negative. However, the overlapping confidence intervals in Figure 1.1 indicate that it does not appear substantially stronger than the effect of immigration facts alone ⁸. The formal tests reported in Appendix 1.A.6 cannot reject that the respondents' reaction to immigration facts is the same, regardless of whether they see a pledge to curb immigration. While this suggests that voters react to factual performance failures, the additional information about an unmet promise does not significantly amplify the effect. This could indicate that voters already expect far-right

⁸Testing this proposition was pre-registered (**Bogatyrev2024b**), together with the four hypotheses outlined in Section 1.2.

governments to curb immigration and penalize them for failing to do so, independent of explicitly reminding them about it in the experiment.

Finally, exposure to negative economic facts provides a useful benchmark. Though its effect points in the negative direction (in line with economic voting), it is losing statistical significance at the 95% confidence level ($p=0.077$ in a regression adjusted for covariates). However, this effect is not statistically distinguishable from the effects of immigration treatments. This suggests that immigration outcomes can be equally, if not more, influential in shaping voter evaluations of far-right governments. Notably, the respondents did not show a strong response to information about the economic slowdown, even though the treatment was informative for the respondents, as the manipulation check shows.

To understand whether the treatments conveyed new information, immediately after the outcome questions, the respondents were asked two questions about recent trends in immigration and economic growth. The manipulation check results (Figure 1.2) indicate that the treatments were indeed informative to a substantial part of the respondents. In particular, providing the immigration facts caused about 20 p.p. increase the probability of correctly choosing that the number of migrant boat arrivals to Italy increased in 2023. Similarly, providing the GDP growth statistics led to 15 p.p. increase in the probability of correctly choosing that the economic growth decreased in 2023. Importantly, these effects were observed despite only 59% of the immigration treatment group and 42% of the economic treatment group confirming in a subsequent question that they trusted the information provided in the vignettes.

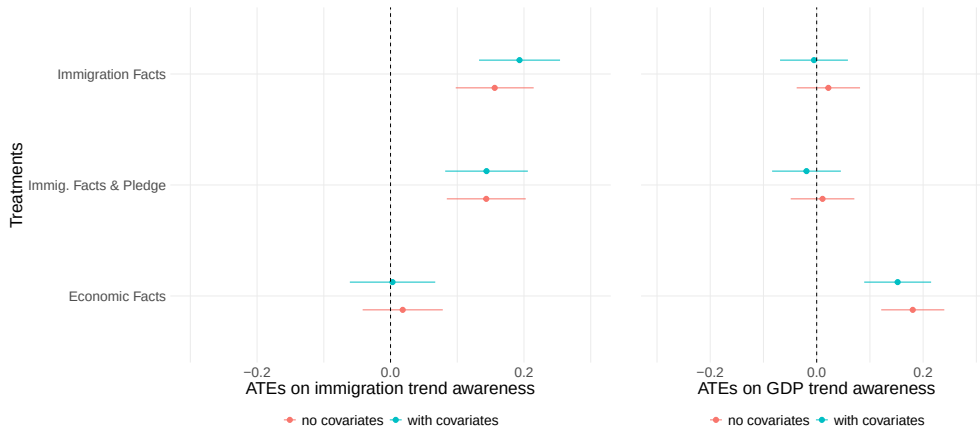


Figure 1.2. Manipulation checks on knowledge about recent trends

Overall, these findings support the main proposition: rising immigration can erode citizens' evaluations of their far-right government and their propensity to vote for it. The next part discusses whether this effect varies substantially across respondents.

1.6.2 Conditional Average Treatment Effects

1.6.2.1 Testing the CATE hypotheses

To gain a deeper understanding of how respondents with particular characteristics vary in their response to rising immigration, I examined the conditional average treatment effects (CATEs). To test whether the effect sizes differ by several key dimensions, I include an interaction term $\beta_5 Immig_i \times Z_i$ into the regression equation:

$$Y_i = \beta_0 + \beta_1 Immig_i + \beta_2 Pled_i + \beta_3 Econ_i + \beta_4 X_i + \beta_5 Immig_i \times Z_i + \beta_6 Z_i + \epsilon_i, \quad (1.2)$$

where Z_i indicates the covariate conditioning the effect: e.g., far-right incumbent voting (hypothesis 2), prioritizing immigration (hypothesis 3), or anti-immigration views (hypothesis 4) for respondent i .

According to hypotheses 2-4,

2. $\beta_5 < 0$ in regressions where Z_i indicates respondents who voted for the incumbent

parties,

3. $\beta_5 < 0$ in regressions where Z_i indicates respondents who consider immigration important,
4. $\beta_5 < 0$ in regressions where Z_i indicates respondents with anti-immigration views.

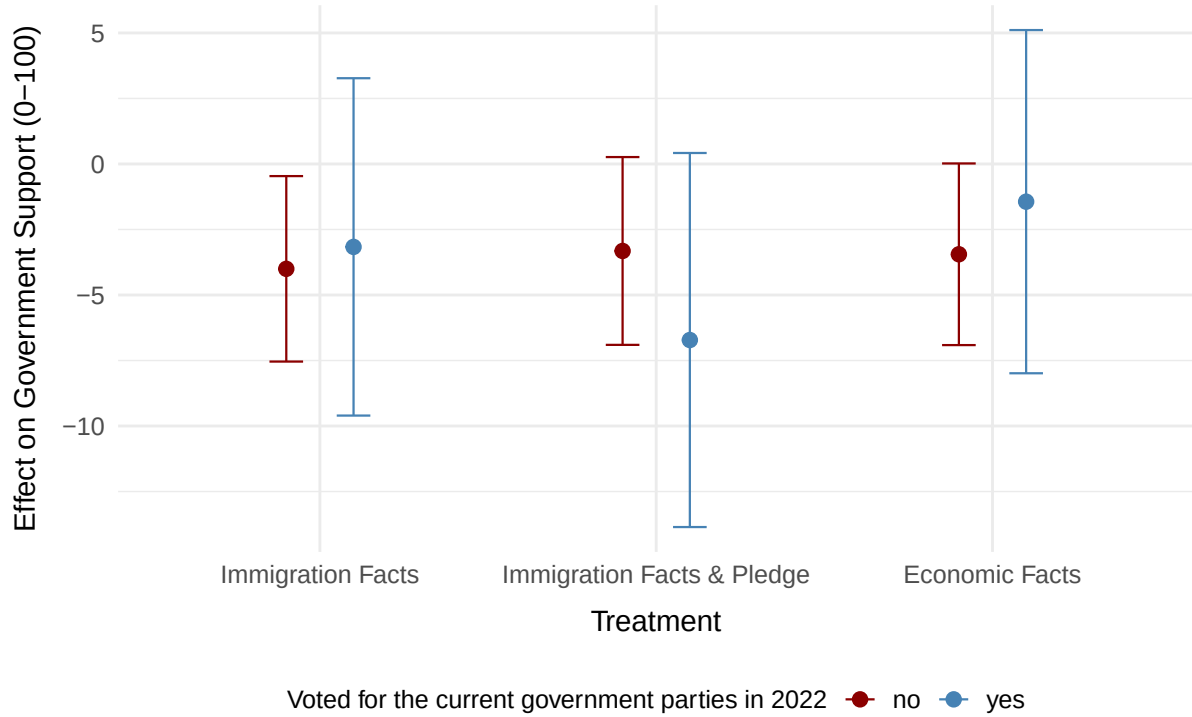
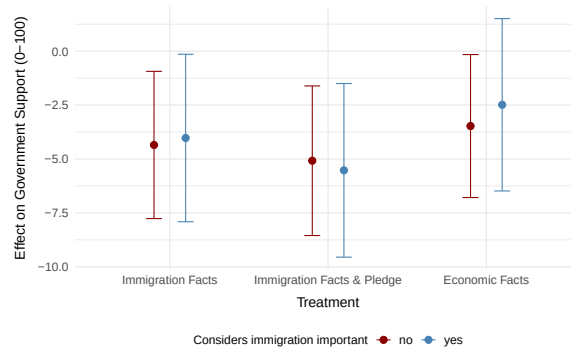


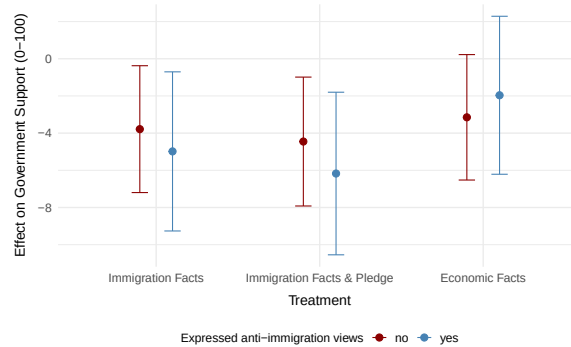
Figure 1.3. Conditional average treatment effects on government support w.r.t. whether the respondent voted for the governing parties

The results indicate that partisan alignment does not significantly moderate accountability for far-right governments support. As shown in Figure 1.3, the estimated treatment effects on government support do not significantly differ between those who voted and who did not vote for the currently governing parties in the last election. The similarity of effects across partisan groups contrasts with expectations from **Hypothesis 2**. This is true both for immigration and economic treatments. This result is also robust to defining partisanship as “having voted for the far right parties” or “having voted for the prime minister’s party”. This suggests that far-right incumbents are not shielded from (nor do

they face increased) accountability pressures among their own supporters—instead. Both co-partisans and non-partisans react similarly to evidence of rising immigration and economic slowdown, reinforcing the idea that, just like the economy, immigration is a salient issue for electoral accountability, even beyond partisan divisions. Furthermore, the results suggest that the immigration treatment effect is not disproportionately driven by the opposition voters, making it unlikely that the observed accountability effect is simply a response to ideological cueing.



(a)



(b)

Figure 1.4. Conditional treatment effects on government support by immigration salience and views

Similarly, the results yield no evidence of heterogeneity in the treatment effects by personal salience of immigration and anti-immigration views. Across all three treatment conditions – exposure to immigration-related factual information, immigration facts combined with a pledge condition, and economic facts – the estimated effects on government

support remain statistically indistinguishable between those who consider immigration an important issue and those who do not (Figure 1.5(a)). Similarly, there is no notable difference in the magnitude of treatment effects between individuals who express anti-immigration attitudes and those who do not (Figure 1.5(b)). Thus, **Hypotheses 3 and 4** are not supported by the data. This suggests that, contrary to the expectations derived from issue voting theory, neither personal salience nor substantive opposition to immigration appears to amplify voters' responsiveness to information about rising immigration when evaluating the far-right government. Instead, the effects of information treatments seem relatively uniform across these subgroups, implying that concerns over immigration are not necessarily activated in a way that conditions accountability in this context.

This is not to say that the levels of personal salience of immigration, anti-immigration views, and prior support for far-right parties were the same among all voters. Figures 1.11, 1.12 and 1.10 show this clearly. While the voters did indeed significantly differ in attitudes toward the government, in every case, the treatment caused effects of similar magnitude in both groups. Both high and low immigration salience voters (Figure 1.11), both voters with and without anti-immigration attitudes (Figure 1.12) and both far-right supporters and others (Figure 1.10) exhibited similar reductions in government approval, perceived competence, and voting intentions. More detailed results of testing the hypotheses about the heterogeneous treatment effects can be found in Appendix 1.A.8.

1.6.2.2 Exploring potential effect heterogeneity

To investigate heterogeneity in treatment effects, I employ the causal forest (CF) methodology (Athey et al., 2019; Wager & Athey, 2018). Unlike standard regression approaches with pre-specified interaction terms, CF is a tree-based ensemble method that recursively partitions the data based on covariate similarities to estimate conditional average treatment effects (CATEs) for subpopulations⁹. This allows CF to flexibly detect and model

⁹Causal forests operate by first separating model estimation into two steps: 1) A treatment assignment model (propensity score estimation); 2) An outcome model (conditional on covariates). After these models are estimated, CF applies residual-on-residual regression, which removes variation due to confounders and isolates the treatment effect in a data-driven manner. Unlike standard decision trees, CF prevents overfitting and allows for valid statistical inference due to honest splitting, meaning that one part of the

treatment effect heterogeneity without strong parametric assumptions about functional form. Furthermore, CF has been shown to strongly outperform saturated OLS regressions in estimating the average and conditional average treatment effects in moderate-to-large samples with a few to a few dozen covariates, as is the case in this study (Fuller & Rametta, 2025). This makes CF a suitable choice for the exploratory analysis.

Using causal forests, I systematically explore the effect heterogeneity with respect to other covariates without having to pre-specify and test each of them individually. In this part, I focus specifically on the immigration facts treatment.

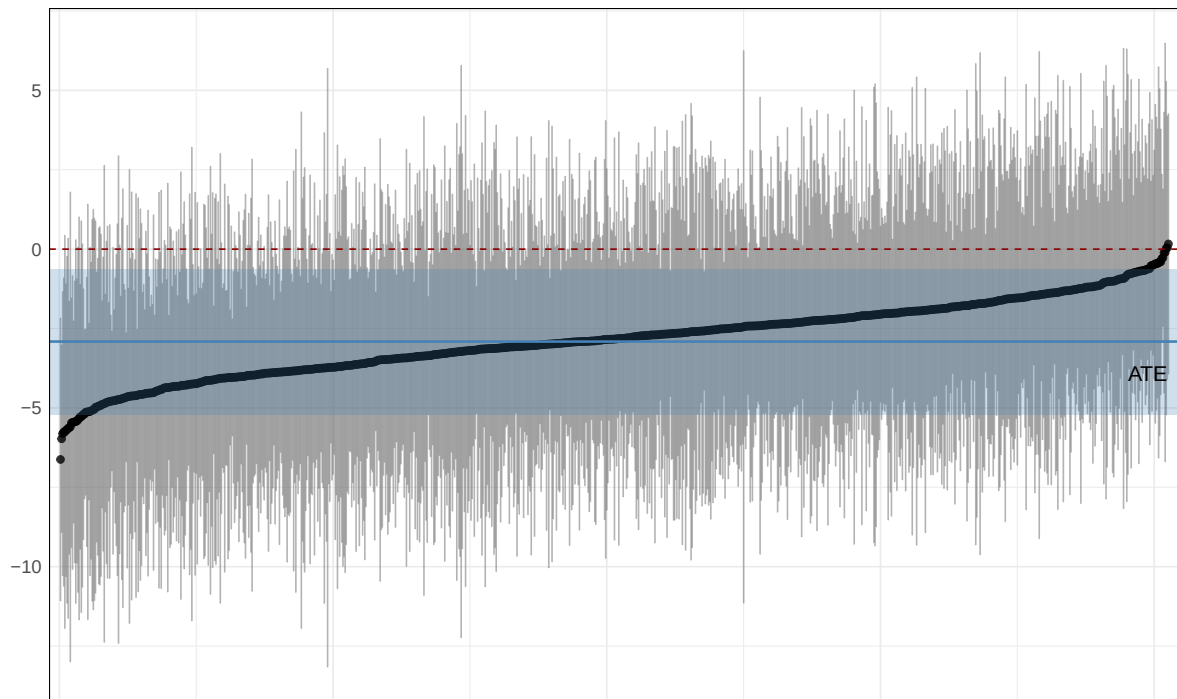


Figure 1.5. *Note:* The blue line represents the point estimate of the ATE. The shaded blue area indicates the 95% confidence interval around the ATE. The gray lines show predicted individual effects. These estimates reflect local average treatment effects for observations with a similar covariate profile rather than precise individual treatment effects (Wager & Athey, 2018).

Using causal forests with 500 trees, I predict individual treatment effects for all data is used to build the partition structure, while another is used for estimation (Athey et al., 2019; Wager & Athey, 2018).

respondents (Figure 1.5). The ATE estimates by the causal forest corroborate my findings from the linear regression in terms of sign and magnitude. Figure 1.5 also shows very little evidence of heterogeneity in treatment effects, as almost none of the individual effects are statistically distinguishable from the average treatment effect. Indeed, 99.8% of the individual confidence intervals include the ATE.

To further inquire into the heterogeneity that is present, I identified the most important covariates associated with it (for details, see Appendix 1.A.8.5). However, none of them showed evidence of moderating heterogeneous treatment effects.

The absence of significant differences suggests that immigration acts as a valence issue – one that is uniformly important across the electorate – rather than a positional issue that polarizes opinions.

1.7 Discussion

This study examines the electoral accountability of far-right governments, focusing on their core political issue in Western advanced economies: immigration. Using a pre-registered survey experiment in Italy, I show that rising immigration can erode support for far-right incumbents. The effect is comparable in magnitude to that of economic performance and does not significantly vary by partisanship, policy preferences, or issue salience. These results suggest that, once in power, far-right parties are subject to performance-based evaluations akin to those found for mainstream incumbents in the literature on electoral accountability (Ashworth, 2012; Healy & Malhotra, 2013; Lewis-Beck & Stegmaier, 2019).

The fact that immigration outcomes shape voter attitudes in a broad-based manner suggests that, rather than being a positional issue that only activates certain segments of the electorate, immigration can function as a valence issue, akin to economic performance. This raises the possibility that other parties could position themselves as its competent managers. In fact, voters in Europe attribute only limited issue competence over immigration to the far right (Abou-Chadi, 2016), and center-right parties are the ones that

arguably “own” the issue in a few European democracies (Kriesi et al., [2024](#)).

Beyond electoral consequences, these results highlight a broader strategic dilemma for far-right incumbents. While, as I show, they are penalized for rising immigration, successfully reducing it may ultimately weaken their electoral mobilization advantage by reducing the salience of the issue. This paradox mirrors the “thermostatic model” (Wlezien, [1995](#)) and the challenges faced by left-wing parties on welfare spending, where delivering on key promises can undermine electoral support (Bernardi & Adams, [2019](#)). Thus, while keeping immigration high, far-right politicians might resort to blame-shifting, agenda-setting, or symbolic politics to mitigate accountability effects like those demonstrated by this study.

While the Italian case provides a valuable setting to examine accountability for the far right, it also presents context-specific factors that may limit generalizability. Italy has experienced substantial immigration in recent decades, and immigration remains a highly salient issue for Italian voters. Moreover, far-right parties differ in their policy platforms and incumbency records across countries. Meloni’s populist far-right majority government that assumed office in 2022 differs from contexts where the far right is a junior coalition partner or faces stronger institutional constraints. Moreover, Italy’s multiparty system allows voters to express discontent by shifting support between multiple alternatives, which may not be the case in two-party systems where accountability mechanisms function differently.

Although experimental evidence provides a strong causal leverage over short-term responses to government performance, long-term electoral consequences remain an open question. However, coming from an experiment conducted just 1.5 years into the government’s term and focusing on immigration outcomes rather than direct policy outputs, my results provide a hard test for accountability. The observed effects likely represent a lower bound of what voters may attribute to the government over a longer tenure or when assessing more direct policy decisions. Moreover, early evaluations are important and can have a lasting impact on voter behavior in the following elections Stiers et al. ([2020](#)).

Future research should examine how accountability dynamics unfold over the electoral cycle. Besides, using observational data may take into account other factors of voting such as media narratives and elite framing.

The accountability patterns demonstrated in this study are compatible with the argument that far-right parties are perceived as ordinary political actors rather than exceptional challengers (Turnbull-Dugarte, [2024](#)). Moreover, they are not invincible. Immigration – long a mobilizing tool for the far right – can also become a liability when they govern. Rather than securing a permanent electoral advantage, the far right’s focus on immigration invites scrutiny, and accountability pressures may ultimately reshape party competition.

References

- About-Chadi, T. (2016). Niche Party Success and Mainstream Party Policy Shifts – How Green and Radical Right Parties Differ in Their Impact. *British Journal of Political Science*, 46(2), 417–436. <https://doi.org/10.1017/S0007123414000155>
- Achen, C. H., & Bartels, L. M. (2016). *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton University Press.
- Adiguzel, F. S., Cansunar, A., & Corekcioglu, G. (2023). Out of Sight, Out of Mind? Electoral Responses to the Proximity of Health Care. *The Journal of Politics*, 85(2), 667–683. <https://doi.org/10.1086/722040>
- Akkerman, T., & de Lange, S. L. (2012). Radical Right Parties in Office: Incumbency Records and the Electoral Cost of Governing. *Government and Opposition*, 47(4), 574–596. <https://doi.org/10.1111/j.1477-7053.2012.01375.x>
- Albertazzi, D., & McDonnell, D. (2015). *Populists in Power*. Routledge. <https://doi.org/10.4324/9781315725789>
- Ansolabehere, S., & Jones, P. E. (2010). Constituents' Responses to Congressional Roll-Call Voting. *American Journal of Political Science*, 54(3), 583–597. <https://doi.org/10.1111/j.1540-5907.2010.00448.x>
- Ashworth, S. (2012). Electoral Accountability: Recent Theoretical and Empirical Work. *Annual Review of Political Science*, 15(1), 183–201. <https://doi.org/10.1146/annurev-polisci-031710-103823>
- Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized Random Forests. *Annals of Statistics*, 47(2), 1148–1178. <https://doi.org/10.1214/18-AOS1709>
- Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and Politics. the Role of Immigration in Shaping Natives' Voting Behavior. *Journal of Public Economics*, 136, 1–13. <https://doi.org/10.1016/j.jpubeco.2016.03.002>
- Bellodi, L., Morelli, M., & Vannoni, M. (2024). A Costly Commitment: Populism, Economic Performance, and the Quality of Bureaucracy. *American Journal of Political Science*, 68(1), 193–209. <https://doi.org/10.1111/ajps.12782>
- Bergmann, J., Hackenesch, C., & Stockemer, D. (2021). Populist Radical Right Parties in Europe: What Impact Do They Have on Development Policy? *JCMS: Journal of Common Market Studies*, 59(1), 37–52. <https://doi.org/10.1111/jcms.13143>
- Bernardi, L., & Adams, J. (2019). Does Government Support Respond to Governments' Social Welfare Rhetoric or Their Spending? An Analysis of Government Support in Britain, Spain and the United States. *British Journal of Political Science*, 49(4), 1407–1429. <https://doi.org/10.1017/S0007123417000199>

- Betz, H.-G., Bernhard, L., & Biard, B. (Eds.). (2019). *Do They Make a Difference? The Policy Influence of Radical Right Populist Parties in Western Europe*. ECPR Press.
- Bischof, D., & Wagner, M. (2019). Do Voters Polarize When Radical Parties Enter Parliament? *American Journal of Political Science*, 63(4), 888–904. <https://doi.org/10.1111/ajps.12449>
- Bisgaard, M. (2019). How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning. *American Journal of Political Science*, 63(4), 824–839. <https://doi.org/10.1111/ajps.12432>
- Bonilla, T. (2022). Promising in Retrospect. In *The Importance of Campaign Promises* (pp. 140–167). Cambridge University Press. <https://doi.org/10.1017/9781108910170.011>
- Bonilla, T. (2024). The Influence of Partisanship on Assessments of Promise Fulfillment and Accountability. *American Political Science Review*, 1–19. <https://doi.org/10.1017/S0003055424000807>
- Braun, D., & Schäfer, C. (2022). Issues That Mobilize Europe. the Role of Key Policy Issues for Voter Turnout in the 2019 European Parliament Election. *European Union Politics*, 23(1), 120–140. <https://doi.org/10.1177/14651165211040337>
- Castanho Silva, B. (2019). “He’s Not One of Them!” Antiestablishment Supporters of Populist Governments in Bolivia and Ecuador. *The Journal of Politics*, 81(3), 1085–1089. <https://doi.org/10.1086/703071>
- Clifford, S., Sheagley, G., & Piston, S. (2021). Increasing Precision Without Altering Treatment Effects: Repeated Measures Designs in Survey Experiments. *American Political Science Review*, 115(3), 1048–1065. <https://doi.org/10.1017/S0003055421000241>
- Cools, S., Finseraas, H., & Rogeberg, O. (2021). Local Immigration and Support for Anti-Immigration Parties: A Meta-Analysis. *American Journal of Political Science*, 65(4), 988–1006. <https://doi.org/10.1111/ajps.12613>
- Cremaschi, S., Rettl, P., Cappelluti, M., & De Vries, C. E. (2024). Geographies of discontent: Public service deprivation and the rise of the far right in Italy. *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12936>
- Dai, Y., & Kustov, A. (2024). The (in)effectiveness of populist rhetoric: A conjoint experiment of campaign messaging. *Political Science Research and Methods*, 12(4), 849–856. <https://doi.org/10.1017/psrm.2023.55>
- De Vries, C. E., & Giger, N. (2014). Holding governments accountable? Individual heterogeneity in performance voting. *European Journal of Political Research*, 53(2), 345–362. <https://doi.org/10.1111/1475-6765.12033>
- De Vries, C. E., & Hobolt, S. B. (2020). *Political Entrepreneurs: The Rise of Challenger Parties in Europe*. Princeton University Press.
- De Vries, C. E., & Solaz, H. (2017). The Electoral Consequences of Corruption. *Annual Review of Political Science*, 20, 391–408. <https://doi.org/10.1146/annurev-polisci-052715-111917>

- Duggan, J., & Martinelli, C. (2017). The Political Economy of Dynamic Elections: Accountability, Commitment, and Responsiveness. *Journal of Economic Literature*, 55(3), 916–984. <https://doi.org/10.1257/jel.20150927>
- Fuller, S., & Rametta, J. T. (2025). *Causal Forest and Doubly Robust Machine Learning for Political Science*. https://doi.org/10.31219/osf.io/nm9de_v2
- Fumarola, A. (2021). Making the Government Accountable: Rethinking Immigration as an Issue in the European Union. *European Politics and Society*, 22(1), 140–159. <https://doi.org/10.1080/23745118.2020.1752453>
- Funke, M., Schularick, M., & Trebesch, C. (2023). Populist Leaders and the Economy. *American Economic Review*, 113(12), 3249–3288. <https://doi.org/10.1257/aer.20202045>
- Gessler, T., Tóth, G., & Wachs, J. (2022). No Country for Asylum Seekers? How Short-Term Exposure to Refugees Influences Attitudes and Voting Behavior in Hungary. *Political Behavior*, 44(4), 1813–1841. <https://doi.org/10.1007/s11109-021-09682-1>
- Golder, M. (2016). Far Right Parties in Europe. *Annual Review of Political Science*, 19(1), 477–497. <https://doi.org/10.1146/annurev-polisci-042814-012441>
- Grzymala-Busse, A., & Nalepa, M. (2022). *How Illiberal Populists Gain and Stay in Power: Programmatic Cohesion and Government Performance* [Unpublished manuscript].
- Hainmueller, J., & Hopkins, D. J. (2014). Public Attitudes Toward Immigration. *Annual Review of Political Science*, 17(1), 225–249. <https://doi.org/10.1146/annurev-polisci-102512-194818>
- Healy, A., & Malhotra, N. (2013). Retrospective Voting Reconsidered. *Annual Review of Political Science*, 16(1), 285–306. <https://doi.org/10.1146/annurev-polisci-032211-212920>
- Healy, A., Persson, M., & Snowberg, E. (2017). Digging into the Pocketbook: Evidence on Economic Voting from Income Registry Data Matched to a Voter Survey. *American Political Science Review*, 111(4), 771–785. <https://doi.org/10.1017/S0003055417000314>
- Hellwig, T. (2014). The Structure of Issue Voting in Postindustrial Democracies. *Sociological Quarterly*, 55(4), 596–624. <https://doi.org/10.1111/tsq.12072>
- Hillygus, D. S., & Shields, T. G. (2008). *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton University Press.
- Ivarsflaten, E. (2008). What Unites Right-Wing Populists in Western Europe?: Re-Examining Grievance Mobilization Models in Seven Successful Cases. *Comparative Political Studies*, 41(1), 3–23. <https://doi.org/10.1177/0010414006294168>
- Jacob, M. S. (2024). Citizens as a Democratic Safeguard? The Sequence of Sanctioning Elite Attacks on Democracy. *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12847>
- Kinder, D. R., & Kam, C. D. (2010). *Us Against Them: Ethnocentric Foundations of American Opinion*. University of Chicago Press.

- Kitschelt, H. (2018). Party Systems and Radical Right-Wing Parties. In J. Rydgren (Ed.), *The Oxford Handbook of the Radical Right*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190274559.013.9>
- Kriesi, H., Altiparmakis, A., Bojár, Á., & Oană, I.-E. (2024). The Electoral Consequences of the Refugee Crisis. In *Coming to Terms with the European Refugee Crisis* (pp. 331–355). Cambridge University Press. <https://doi.org/10.1017/9781009456555.019>
- Kriesi, H., Grande, E., Dolezal, M., Helbling, M., Höglinger, D., Hutter, S., & Wüest, B. (2012). *Political Conflict in Western Europe*. Cambridge University Press.
- Kriner, D. L., & Shen, F. X. (2010). *The Casualty Gap: The Causes and Consequences of American Wartime Inequalities*. Oxford University Press.
- Krishnarajan, S., & Jensen, C. (2022). When Is a Pledge a Pledge? *British Journal of Political Science*, 52(4), 1911–1922. <https://doi.org/10.1017/S0007123421000284>
- Kustov, A., & Landgrave, M. (2023). *Immigration Is Difficult?! Informing Voters About Immigration Policy Fosters Pro-Immigration Views*. <https://doi.org/10.31219/osf.io/mu4j5>
- Lewis-Beck, M. S., & Stegmaier, M. (2019). Economic Voting. In R. D. Congleton, B. Grofman, & S. Voigt (Eds.), *The Oxford Handbook of Public Choice*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190469733.013.12>
- Lewis-Beck, M. S., & Nadeau, R. (2011). Economic voting theory: Testing new dimensions. *Electoral Studies*, 30(2), 288–294. <https://doi.org/10.1016/j.electstud.2010.09.001>
- lo Polito, F. (2023). *Italians First: Refugee Reception Policy and National Identity* [Unpublished manuscript]. <https://gsipe-workshop.github.io/files/JMP%20-%20Federica%20lo%20Polito.pdf>
- Lutz, P. (2019). Variation in Policy Success: Radical Right Populism and Migration Policy. *West European Politics*, 42(3), 517–544. <https://doi.org/10.1080/01402382.2018.1504509>
- Margalit, Y. (2019). Political Responses to Economic Shocks. *Annual Review of Political Science*, 22(1), 277–295. <https://doi.org/10.1146/annurev-polisci-050517-110713>
- Matthieß, T. (2020). Retrospective pledge voting: A comparative study of the electoral consequences of government parties' pledge fulfilment. *European Journal of Political Research*, 59(4), 774–796. <https://doi.org/10.1111/1475-6765.12377>
- Matthieß, T. (2022). Retrospective pledge voting and mistrusting citizens: Evidence for the electoral punishment of pledge breakage from a survey experiment. *Electoral Studies*, 80, 102547–102547. <https://doi.org/10.1016/j.electstud.2022.102547>
- Matthieß, T. (2024). *Retrospective Pledge Voting and Political Accountability: The Costs of Broken Promises*. Springer Nature.
- Mudde, C. (2013). Three Decades of Populist Radical Right Parties in Western Europe: So What? *European Journal of Political Research*, 52(1), 1–19. <https://doi.org/10.1111/j.1475-6765.2012.02065.x>
- Mudde, C. (2019). *The Far Right Today*. Polity.

- Nadeau, R., Lewis-Beck, M. S., & Bélanger, É. (2013). Economics and Elections Revisited. *Comparative Political Studies*, 46(5), 551–573. <https://doi.org/10.1177/0010414012463877>
- Naurin, E., Soroka, S., & Markwat, N. (2019). Asymmetric accountability: An experimental investigation of biases in evaluations of governments' election pledges. *Comparative Political Studies*, 52(13–14), 2207–2234. <https://doi.org/10.1177/0010414019830740>
- Naurin, E., & Thomson, R. (2020). The Fulfilment of Election Pledges. In *Research Handbook on Political Representation* (pp. 289–300). Edward Elgar Publishing.
- Neundorff, A., & Adams, J. (2018). The Micro-Foundations of Party Competition and Issue Ownership: The Reciprocal Effects of Citizens' Issue Salience and Party Attachments. *British Journal of Political Science*, 48(2), 385–406. <https://doi.org/10.1017/S0007123415000642>
- Paxton, F. (2023). *Restrained Radicals: Populist Radical Right Parties in Local Government*. Cambridge University Press. <https://doi.org/10.1017/9781009379069>
- Pirro, A. L. P. (2023). Far right: The significance of an umbrella concept. *Nations and Nationalism*, 29(1), 101–112. <https://doi.org/10.1111/nana.12860>
- Porter, E., & Velez, Y. R. (2022). Placebo Selection in Survey Experiments: An Agnostic Approach. *Political Analysis*, 30(4), 481–494. <https://doi.org/10.1017/pan.2021.16>
- Rooduijn, M., Pirro, A. L. P., Halikiopoulou, D., Froio, C., Kessel, S. V., Lange, S. L. D., Mudde, C., & Taggart, P. (2023). The PopuList: A database of populist, far-left, and far-right parties using expert-informed qualitative comparative classification (EiQCC). *British Journal of Political Science*, 1–10. <https://doi.org/10.1017/S0007123423000431>
- Rose, R., & Wessels, B. (2019). Money, Sex and Broken Promises: Politicians' Bad Behaviour Reduces Trust. *Parliamentary Affairs*, 72(3), 481–500. <https://doi.org/10.1093/pa/gsy024>
- Royed, T. J., Leyden, K. M., & Borrelli, S. A. (2000). Is 'Clarity of Responsibility' Important for Economic Voting? Revisiting Powell and Whitten's Hypothesis. *British Journal of Political Science*, 30(4), 669–698. <https://doi.org/10.1017/S0007123400210296>
- Steinberg, D. A. (2022). How Voters Respond to Currency Crises: Evidence from Turkey. *Comparative Political Studies*, 55(8), 1332–1365. <https://doi.org/10.1177/00104140211060268>
- Stiers, D., Dassonneville, R., & Lewis-Beck, M. S. (2020). The Abiding Voter: The Lengthy Horizon of Retrospective Evaluations. *European Journal of Political Research*, 59(3), 646–668. <https://doi.org/10.1111/1475-6765.12360>
- Tertychnaya, K., De Vries, C. E., Solaz, H., & Doyle, D. (2018). When the Money Stops: Fluctuations in Financial Remittances and Incumbent Approval in Central Eastern Europe, the Caucasus and Central Asia. *American Political Science Review*, 112(4), 758–774. <https://doi.org/10.1017/S0003055418000485>

-
- Thompson, D. M. (2023). Do Citizens Vote Against Incumbents Who Permit Local Immigration? Evidence from the Mariel Boatlift. *Political Science Research and Methods*, 11(4), 728–745. <https://doi.org/10.1017/psrm.2021.76>
- Tolsma, J., Laméris, J., & Savelkoul, M. (2021). Exposure to asylum seekers and changing support for the radical right: A natural experiment in the Netherlands. *PLOS One*, 16(2), e0245644. <https://doi.org/10.1371/journal.pone.0245644>
- Turnbull-Dugarte, S. J. (2024). Public Support for the Cordon Sanitaire: Descriptive Evidence from Spain. *Party Politics*, 13540688241246141. <https://doi.org/10.1177/13540688241246141>
- Valentim, V. (2021). Parliamentary Representation and the Normalization of Radical Right Support. *Comparative Political Studies*, 54(14), 2475–2511. <https://doi.org/10.1177/0010414021997159>
- van de Wardt, M., De Vries, C. E., & Hobolt, S. B. (2014). Exploiting the Cracks: Wedge Issues in Multiparty Competition. *The Journal of Politics*, 76(4), 986–999. <https://doi.org/10.1017/S0022381614000565>
- Visconti, G. (2019). Economic Perceptions and Electoral Choices: A Design-Based Approach. *Political Science Research and Methods*, 7(4), 795–813. <https://doi.org/10.1017/psrm.2017.26>
- Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523), 1228–1242. <https://doi.org/10.1080/01621459.2017.1319839>
- Weldon, S., & McNeney, D. (2019). Issue Voting: Modern and Classic Accounts. In *Oxford Research Encyclopedia of Politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.776>
- Wlezien, C. (1995). The Public as Thermostat: Dynamics of Preferences for Spending. *American Journal of Political Science*, 39(4), 981–1000. <https://doi.org/10.2307/2111666>

1.A Appendix

1.A.1 Covariate Coding

Specifically, the variable *female* indicates the respondent's gender, coded as 1 for females and 0 for males. *Age* is measured in years. *Education* (*educ*) is quantified in total years of schooling, ranging from elementary education (5 years) to doctoral level (20 years). *Immigration background* (*migr*) is a binary variable indicating whether the respondent is foreign-born.

Regional variables include *reg_north* and *reg_center*, indicating whether the respondent resides in the northern or central regions of Italy, respectively. Marital status (*married*) and parental status (*parent*) are binary variables indicating whether the respondent is married and whether they have children. *Employed* is a binary variable indicating the respondent's employment status.

Income (*inc*) is recoded into a continuous variable representing annual income levels. *Financial support* (*fin_supp*) is a binary variable indicating whether the respondent received financial help from others. *Financial gain* (*fin_gain*) measures the change in personal finances over the past year. *Financial stress* (*fin_stress*) indicates the difficulty in managing the personal budget, and *natural hazards* (*natur_haz*) indicates whether the respondent has suffered from a natural hazard.

Anti-feminist attitudes (*i_nofem*) and anti-LGBTQ attitudes (*i_nolgbt*) are indexes generated using principal component analysis (PCA) from responses to relevant survey questions.

Additionally, political attitudes include voting for far-right parties in the 2022 elections (*fr_2022*), anti-immigration attitudes (*anti_imm*), and the importance of immigration issues (*mig_imp_bin*). The variable *att_check* indicates whether the respondent passed an attention check.

1.A.2 Covariate Balance Assessment

Before estimating the effect in the vignette experiment, I first test whether randomization was successful by regressing socioeconomic variables (gender, age, immigration background, education, civil and parental status, employment, receiving financial support and recent changes in personal finance) on exposure to the different vignettes. Successful randomization implies that none of these characteristics have a statistically significant association with exposure to the different vignettes. This is indeed the case. The results of assessment are presented in Table ??.

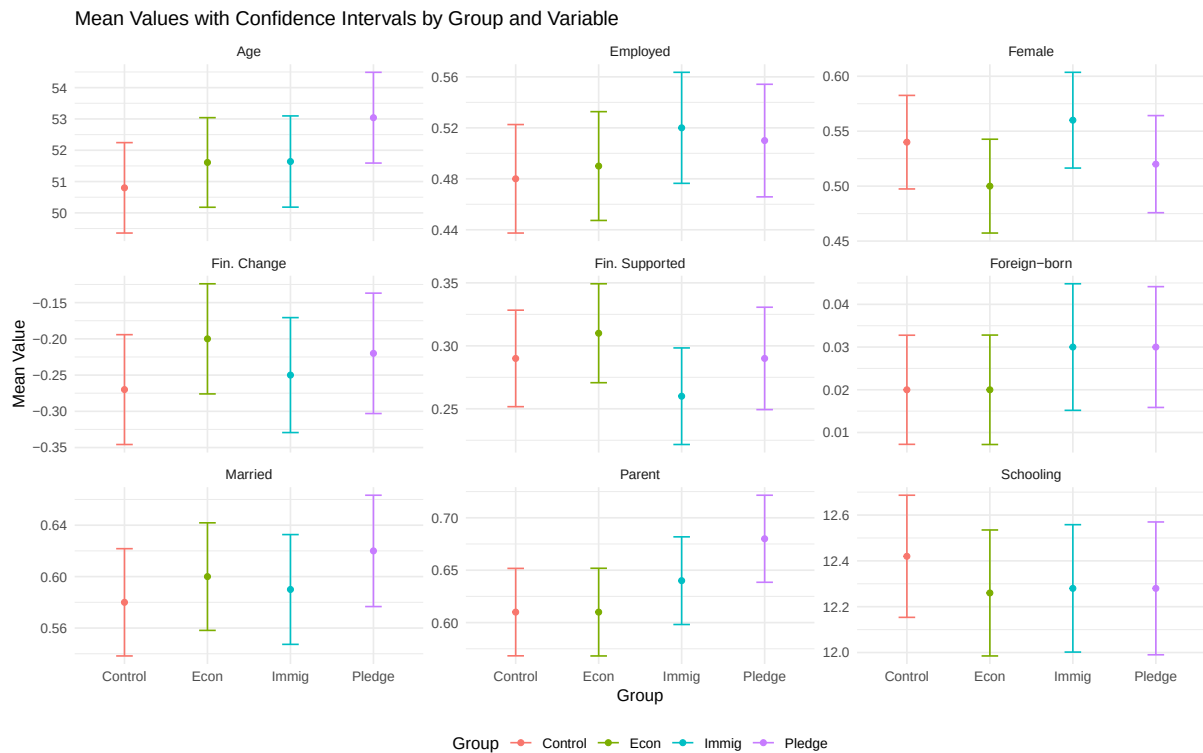


Figure 1.6. Average values of nine socioeconomic variables for each experimental group

I also compare the averages of these key socioeconomic variables among all groups (Figure 1.6). Figure 1.6 displays the average values of nine key socioeconomic variables for each of the four experimental groups. Each subplot in the figure corresponds to one of these variables and shows the mean value for each experimental group, along with the

confidence intervals. The vertical bars represent the confidence intervals around the mean values. The proportion of females in each group is approximately around 0.5, indicating a roughly equal distribution across groups. The age variable shows mean ages around the low 50s for all groups, with no significant deviations, while the foreign-born variable has very low means around 0.02-0.03 across all groups. The mean years of schooling are slightly above 12 for all groups, reflecting the varied educational attainment of the participants. The recent changes in personal finance present small negative changes, with values between -0.35 and -0.15. The financial support variable indicates that around 0.25-0.35 of the participants received financial support from others. The marital status variable shows that the proportion of married participants ranges from 0.58 to 0.62 across the groups. The parental status variable shows that the proportion of participants who have children ranges from 0.61 to 0.68. Finally, the employment status variable shows that the proportion of employed participants is around 0.48-0.56 across groups. These results confirm that the randomization was successful, as there are no significant differences in these key socioeconomic variables across the experimental groups.

	control (N=530)	econ (N=527)	immig (N=506)	pled (N=492)	p value	F-Stat (p.value)
female					0.345	1.11 (0.34)
Mean (SD)	0.54 (0.50)	0.50 (0.50)	0.56 (0.50)	0.52 (0.50)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
age					0.195	1.57 (0.2)
Mean (SD)	50.80 (16.96)	51.61 (16.76)	51.64 (16.72)	53.04 (16.40)		
Range	18.00 - 94.00	18.00 - 90.00	18.00 - 98.00	18.00 - 92.00		
foreign-born					0.912	0.28 (0.84)
Mean (SD)	0.02 (0.15)	0.02 (0.15)	0.03 (0.17)	0.03 (0.16)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
schooling					0.842	0.18 (0.91)
Mean (SD)	12.42 (3.13)	12.26 (3.22)	12.28 (3.19)	12.28 (3.28)		
Range	5.00 - 20.00	5.00 - 20.00	5.00 - 20.00	5.00 - 20.00		
married					0.504	0.78 (0.5)
Mean (SD)	0.58 (0.49)	0.60 (0.49)	0.59 (0.49)	0.62 (0.49)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
parent					0.082 [†]	2.24 (0.08)
Mean (SD)	0.61 (0.49)	0.61 (0.49)	0.64 (0.48)	0.68 (0.47)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
employed					0.538	0.72 (0.54)
Mean (SD)	0.48 (0.50)	0.49 (0.50)	0.52 (0.50)	0.51 (0.50)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
fin. supported					0.497	0.79 (0.5)
Mean (SD)	0.29 (0.45)	0.31 (0.46)	0.26 (0.44)	0.29 (0.46)		
Range	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00	0.00 - 1.00		
fin. change					0.592	1 (1)
Mean (SD)	-0.27 (0.89)	-0.20 (0.89)	-0.25 (0.91)	-0.22 (0.94)		
Range	-2.00 - 2.00	-2.00 - 2.00	-2.00 - 2.00	-2.00 - 2.00		

[†] p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

1.A.3 Post-Treatment Questions

1.A.3.1 Outcome Questions

After the experiment, I ask all participants the same outcome questions about their perceptions of the Italian government and its head. Like the entire survey, the outcome questions are provided to the respondents in Italian. The following English version is the closest translation from Italian to be used in this paper and later analyses.

The questions are preceded by a preface that reminds the participants when the current government took office and what parties are in it:

*“Since October 2022, the Italian government is headed by Georgia Meloni.
The government includes three political parties: Brothers of Italy, Lega and
Forza Italia. We will now ask you a few questions on your opinions about the
government.”*

To account for ordering effects, I randomize the order in which participants are asked the first three questions. The open-ended question always comes fourth.

1. Approval

“To what extent do you approve of the Italian government?”

Eleven ordered response options, from 0 (*not at all*) to 10 (*a lot*), as well as “*prefer not to respond*”.

2. Competence

*“To what extent do you agree that the Italian government is competent
for their job?”*

Five ordered response options from “*strongly agree*” to “*strongly disagree*” plus “*prefer not to respond*”.

3. Voting propensity

“If the elections were tomorrow, how likely are you to vote for any of the government parties?”

Eleven ordered response options, from *0 (not at all)* to *10 (a lot)*, as well as *“prefer not to respond”*.

1.A.3.2 Knowledge and Trust Checks

After the outcome questions, I ask all (both treated and control) respondents what they believed about the trends in immigration and economic growth. Asking this is useful to understand whether the treatment has affected anyone’s beliefs about immigration and economic growth. Besides, I ask all treated (not control) respondents whether they believed the statistical information in the vignettes. This question is aimed to empirically assess whether respondents receive and perceive the treatment as intended.

I ask about beliefs after the close-ended outcomes to avoid priming the respondents to think about trends they were not exposed to in the vignettes. I also ask about trust after the close-ended outcomes to avoid priming the treated respondents to doubt the information we have provided. However, I put both checks before the last question to minimize non-response due to survey fatigue, which might be especially high due to writing an open-ended comment.

Like the entire survey, these questions are provided to the respondents in Italian. The following English version is the closest literal translation from Italian to be used in this paper and in later analyses.

- Knowledge: economic trend

“According to you, how did the economic growth in Italy change in 2023?”

Response options: *accelerated, did not change, decreased, I don’t know.*

- Knowledge: immigration trend

“According to you, how did the number of migrant boat arrivals to Italy change in 2023?”

Response options: *increased, did not change, decreased, I don’t know.*

- Trust in the provided information¹⁰

“A few questions ago, we have shown you a graph with statistical information. Do you trust the information provided in the graph?”

Three response options: *yes, no, I don't know*

¹⁰This question is not shown to the control group, who did not see any vignette.

1.A.4 Creating an index of government support

Table 1.2 presents the results of the principle component analysis (PCA) of the three outcome variables: government approval, perceived government competence and propensity to vote for the government. The high and similar loadings of PC1 across all three variables demonstrate that the first principle component captures a common dimension among all three outcomes, thus supporting its usage as an index.

	PC1	PC2	PC3
Standard deviation	1.598	0.531	0.405
Proportion of Variance	0.851	0.094	0.055
Cumulative Variance	0.851	0.945	1.000
Variable Loadings			
Government Approval	0.591	0.014	0.806
Government Competence	0.571	0.699	-0.431
Government Vote Propensity	0.570	-0.715	-0.405

Table 1.2. Principal Component Analysis of the Outcome Variables

Additionally, Table 1.3 shows the correlation between the government support index created via PCA and the original variables, with all correlations with the index exceeding 0.91, indicating a strong summary measure. Besides, the table demonstrates that the three original measures all also strongly correlated among each other.

	Gov Approval	Gov Competence	Gov Vote	PC1
Gov Approval	1.000	0.808	0.804	0.945
Gov Competence	0.808	1.000	0.718	0.912
Gov Vote	0.804	0.718	1.000	0.910
Gov Support (PC1)	0.945	0.912	0.910	1.000

Table 1.3. Correlation Matrix of Original Variables and PC1

1.A.5 Main results disaggregated by outcome items

My theoretical framework predicts that exposure to information about increasing immigration will negatively affect citizens' attitudes toward a far-right government (hypothesis 1). I test the hypotheses using every outcome question, estimating the effects of the treatments on individual responses about different outcomes Y_i :

- (a) government approval
- (b) government competence
- (c) propensity to vote for the government

The experimental results support this hypothesis for all three attitudes analyzed. Figure 1.7 and Table 1.4 present the average treatment effects (ATEs) of the experimental vignettes relative to the control group, estimated via ordinary least squares (OLS) regression. The baseline specification includes covariates listed in the Variables section because controlling for them improves precision (Clifford et al., 2021). As Figure 1.7 shows, the inclusion of covariates narrowsthe confidence intervals without substantially changing the point estimates of the treatment effects.

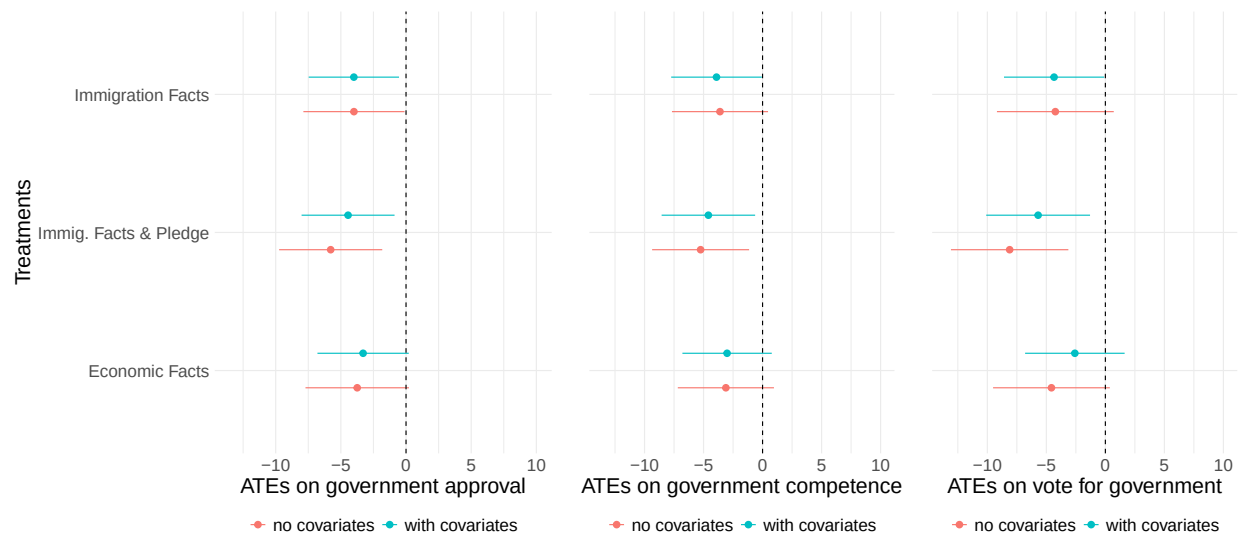


Figure 1.7. Average treatment effects of each experimental vignette relative to the control group

The ATEs indicate that exposure to information about rising immigration signifi-

cantly decreases approval, perceived competence, and voting intentions for the far-right government. Specifically, the ‘Immigration Facts’ treatment led to significant ($p < 0.05$) decreases across all three dimensions: government approval by 4.0 percentage points, perceived competence by 3.9 percentage points, and voting propensity by 4.3 percentage points. The similarity in the magnitude of these effects suggests that disapproval, perceptions of incompetence, and decreased voting intentions are closely interrelated responses to immigration information.

	Approval	Competence	Vote Intention
Immigration Facts	−4.0* (1.8)	−3.9* (2.0)	−4.3* (2.2)
Immig. Facts & Pledge	−4.4* (1.8)	−4.6* (2.0)	−5.7* (2.2)
Economic Facts	−3.3† (1.8)	−3.0 (1.9)	−2.6 (2.1)
Num.Obs.	1616	1625	1572
Covariates	Yes	Yes	Yes

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Covariates included: gender, age, education, macro-region, migration background, marital status, parental status, employment status, income, financial gains, financial support, financial stress, natural hazard exposure, feminist attitudes, LGBTQ attitudes, anti-immigration attitudes, far-right vote in 2022, attention check.

Table 1.4. Estimates of Treatment Effects on Approval, Perceived Competence, and Vote Intention

The OLS estimates are further corroborated by the causal forest (500 trees) estimates with similar magnitude, shown in Figure 1.8 for the immigration treatment. Among the individual confidence intervals, 98.6% include the ATE for government approval, 99.5% include the ATE for government competence and 99.8% include the ATE for electoral support of government parties.

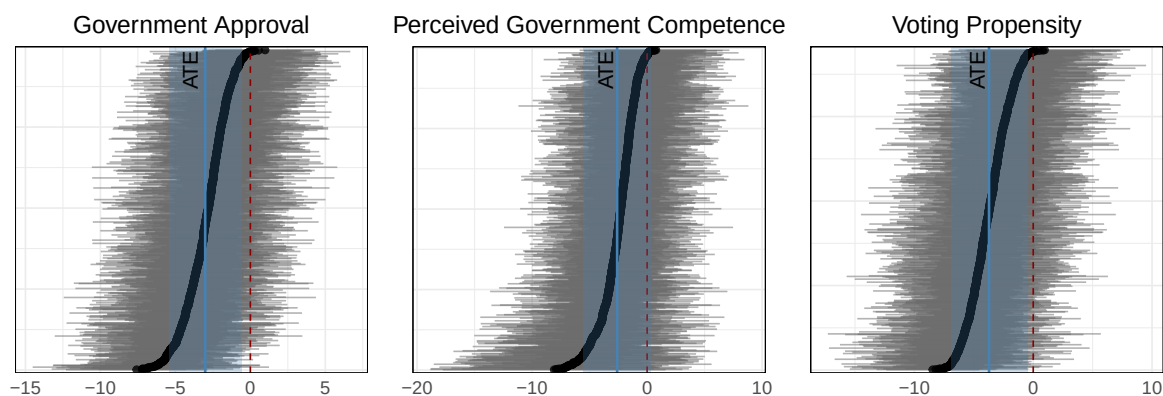


Figure 1.8. *Note:* The gray lines show predicted individual effects. The blue lines represent the point estimates of the average treatment effect (ATE). The shaded blue areas indicate the 95% confidence interval around the ATE.

For the outcome-disaggregated results on conditional treatment effects, see Appendix [1.A.8](#).

1.A.6 Testing the equivalence of Vignettes 1 and 2

The ATE estimates are based on the following regression:

$$Y_i = \beta_0 + \beta_1 Immig_i + \beta_2 Pled_i + \beta_3 Econ_i + \beta_4 X_i + \epsilon_i, \quad (1.3)$$

where for respondent i :

- β_1 is a coefficient measuring the effect of treatment 1 (immigration facts) on the outcome Y_i ,
- β_2 is a coefficient measuring the effect of treatment 2 (immigration facts and pledge) on the outcome Y_i .

For every outcome Y (government approval, government competence and propensity to vote for the government), I am testing whether the respondents' reaction to immigration facts is the same whether or not they see a pledge to curb immigration, i.e. whether $|\beta_1| = |\beta_2|$.

	Wald-stat	Std. Error	z	Pr(> z)
Approval	0.444	1.81	0.246	0.806
Competence	0.688	2.04	0.337	0.736
Vote Intention	1.35	2.21	0.612	0.541

Note: The table shows the results of three linear hypothesis tests comparing the effects of the *Immigration Facts* and *Immigration Facts & Pledge* treatments on government approval, competence, and voting intention. The tests were conducted using the `hypotheses` function from the `marginalEffects` package in R.

Table 1.5. Wald tests for equivalence of the vignettes 1 and 2

The Wald tests shown in Table 1.5 indicate that for all three outcomes – government approval, competence, and vote intention – the differences between these two treatments are statistically insignificant (with p-values of 0.81, 0.74, and 0.54, respectively). Thus,

respondent reactions to solely immigration facts (treatment 1) are statistically equivalent to their reactions when receiving immigration facts combined with a pledge to curb immigration (treatment 2). This suggests that the pledge to curb immigration does not significantly alter the effects already produced by providing immigration facts alone. Potentially, this could be explained by the high salience of Meloni's anti-immigration platform, i.e. both groups already knowing what Meloni's commitments were.

1.A.7 Sentiment towards the incumbent

At the end of the survey, participants were asked to share their opinions on government performance in a short commentary that would be anonymized and delivered to the Italian Prime Minister. 1152 out of 2055 respondents agreed to do this, and 1039 of them wrote responses between 10 and 300 characters (maximum length). For the analysis, I translated these comments from Italian into English via Google Translate.

Using a transformer model BART, I assigned a score between 0 and 1 on several sentiment dimensions to each of these responses. I then regressed each of these sentiment scores on the treatment indicators, similarly to estimating the average treatment effects on the main outcome variables. The 95% confidence intervals for these effects almost always include zero, though some of the coefficients are marginally significant at this or larger levels of alpha. Some coefficients seem consistent in point estimate direction for different treatments (e.g. skeptical, frustrated, distrustful, disappointed, demanding and angry). Though this analysis does not provide clear evidence for any effects, perhaps the statistical power is too small to catch these effects. The power of this analysis is lower than the main analysis due to only observing about half of the respondents writing the comments. Admittedly, any means of automated text classification, even as advanced as transformer models, also do not provide perfect outcome measures.

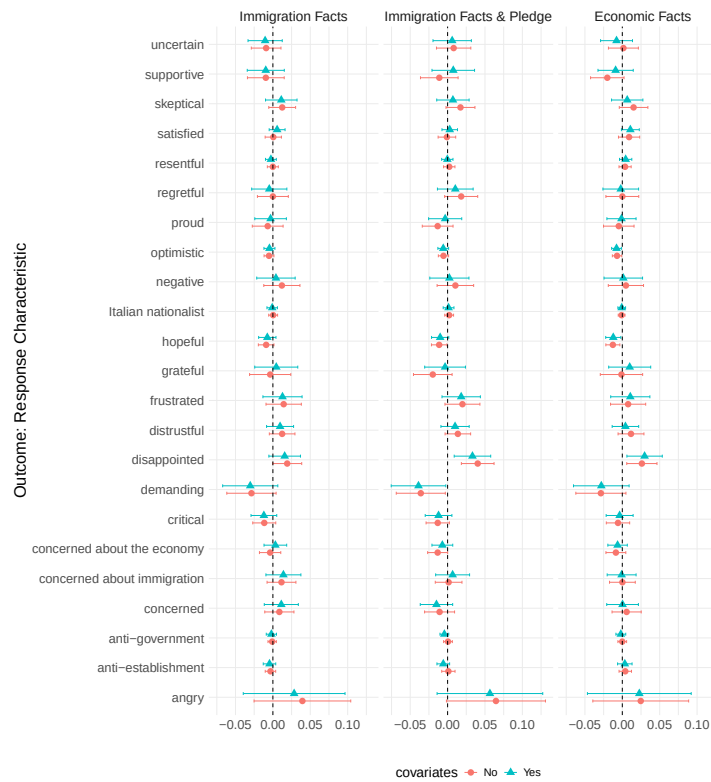


Figure 1.9. Average treatment effects on the sentiment of comments addressed to the prime minister

1.A.8 Conditional Average Treatment Effects

1.A.8.1 Conditional Treatment Effects - Predictions

For all outcome variables – government approval, perceived competence and electoral support – the effects do not differ conditional on these three measures of voter ideology. In particular, the effect of rising immigration was not moderated by these variables, and the hypotheses 2, 3 and 4 stated above are rejected. The detailed estimation results are provided further.

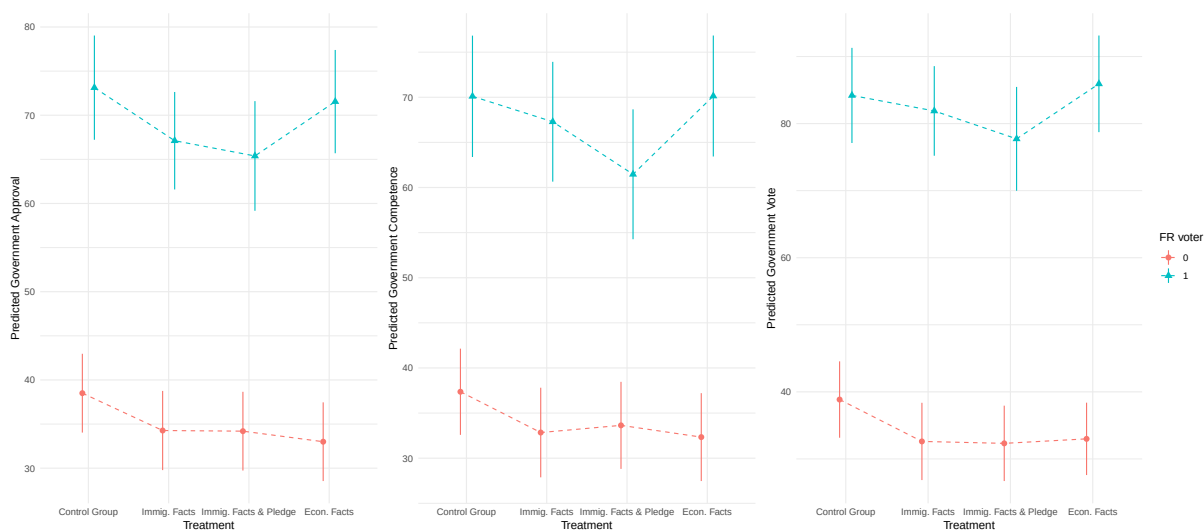


Figure 1.10. Predicted values of outcome variables conditional on far-right voting

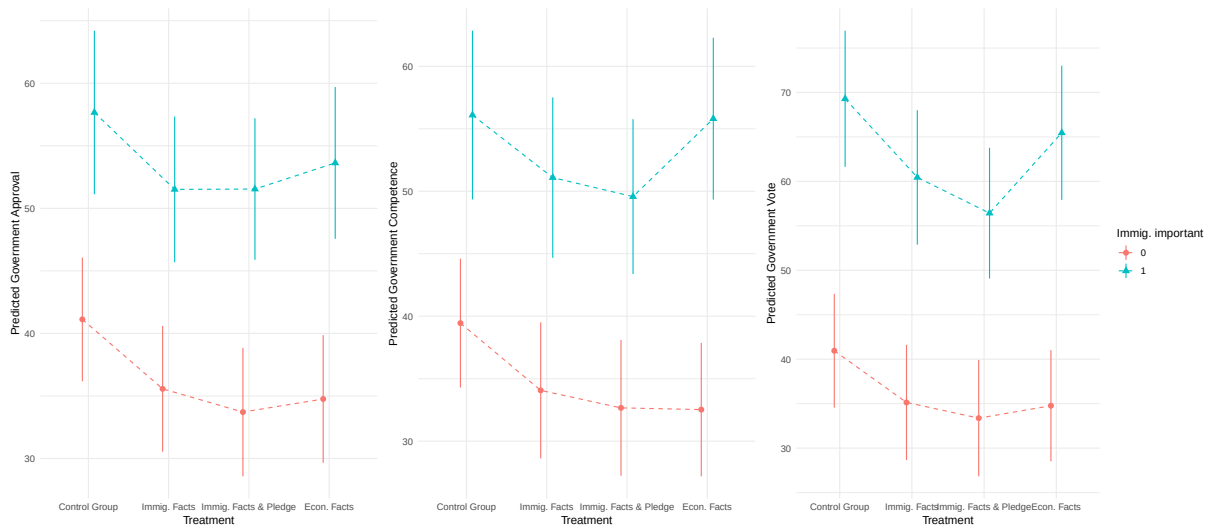


Figure 1.11. Predicted values of outcome variables conditional on the salience of immigration

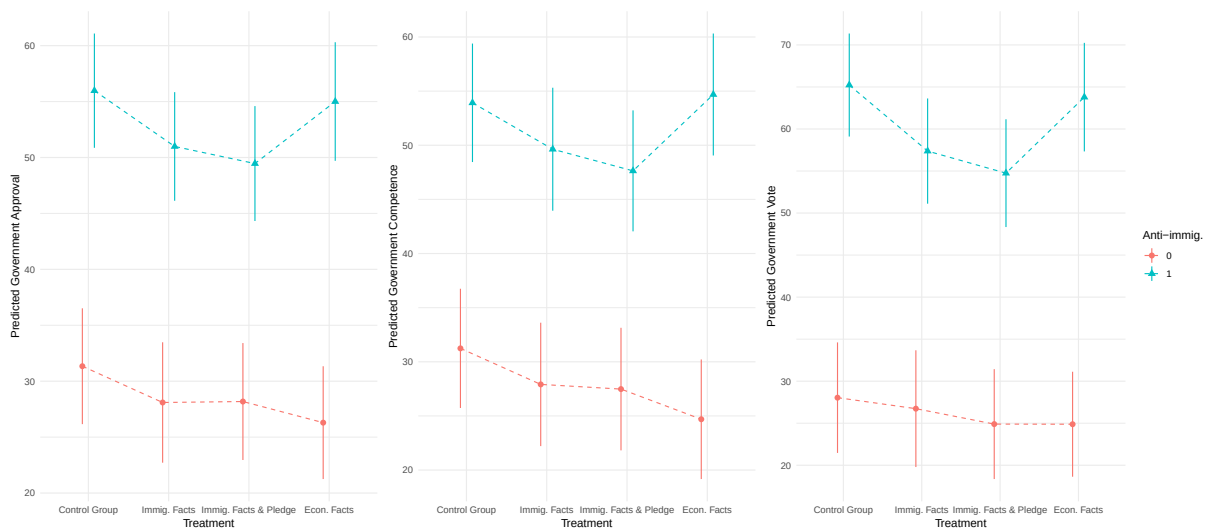


Figure 1.12. Predicted values of outcome variables conditional on anti-immigration views

1.A.8.2 Conditional Average Treatment Effects by Salience of Immigration

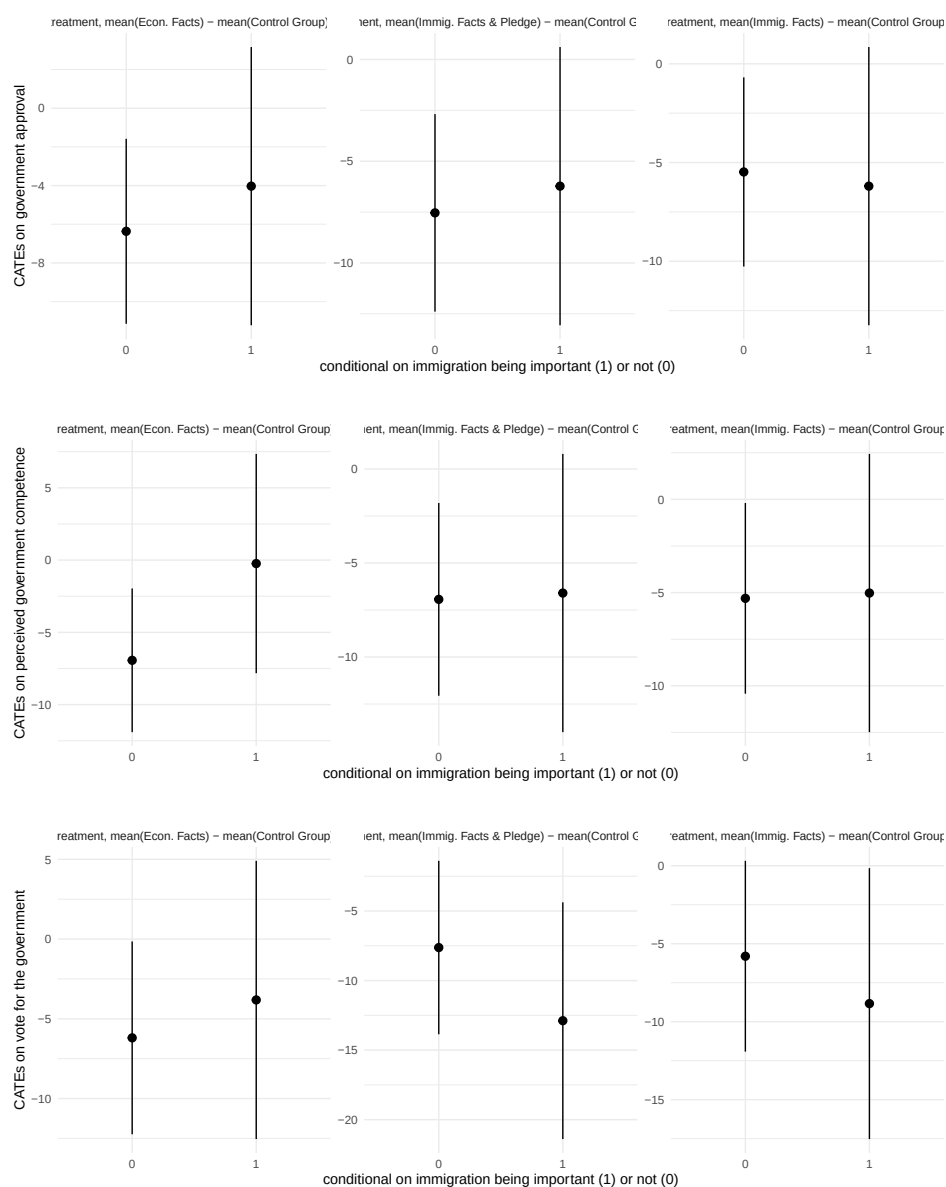


Figure 1.13. Conditional average treatment effects on attitudes to the government w.r.t. whether the respondent voted considered immigration important

	Approval	Competence	Vote Intention
Immig. facts	-5.5* (2.4)	-5.3* (2.6)	-5.8† (3.1)
Immig. facts & pledge	-7.5** (2.5)	-6.9** (2.6)	-7.6* (3.2)
Econ. facts	-6.4** (2.4)	-6.9** (2.5)	-6.2* (3.1)
Immig. important	16.6*** (3.2)	16.7*** (3.3)	28.4*** (3.9)
Immig. facts x Immig. important	-0.7 (4.4)	0.3 (4.6)	-3.0 (5.4)
Immig. facts & pledge x Immig. important	1.3 (4.3)	0.3 (4.6)	-5.3 (5.4)
Econ. facts x Immig. important	2.3 (4.4)	6.7 (4.6)	2.4 (5.4)
Num.Obs.	1649	1658	1604
Covariates	Yes	Yes	Yes

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.6. Estimates of conditional average treatment effects via OLS interaction terms with importance of immigration

1.A.8.3 Conditional Average Treatment Effects by Anti-Immigration Views

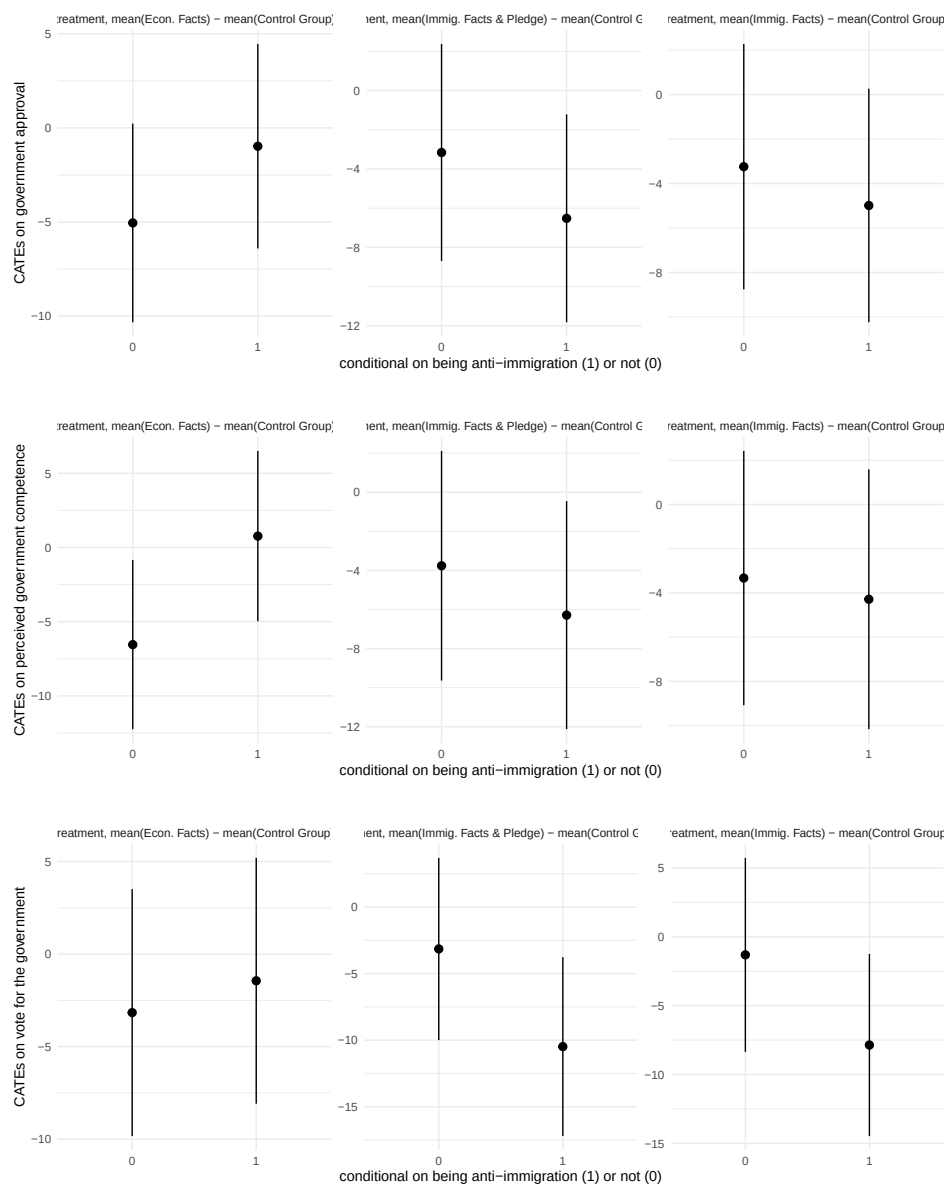


Figure 1.14. Conditional average treatment effects on attitudes to the government w.r.t. whether the respondent expressed anti-immigration attitudes

	Approval	Competence	Vote Intention
Immig. facts	−3.2 (2.8)	−3.3 (2.9)	−1.3 (3.6)
Immig. facts & pledge	−3.2 (2.8)	−3.8 (3.0)	−3.1 (3.5)
Econ. facts	−5.0† (2.7)	−6.5* (2.9)	−3.2 (3.4)
Anti-immig.	24.6*** (2.8)	22.7*** (2.9)	37.2*** (3.5)
Immig. facts x Anti-immig.	−1.7 (3.9)	−1.0 (4.2)	−6.6 (5.0)
Immig. facts & pledge x Anti-immig.	−3.4 (3.9)	−2.5 (4.2)	−7.3 (4.9)
Econ. facts x Anti-immig.	4.1 (3.9)	7.3† (4.1)	1.7 (4.8)
Num.Obs.	1616	1625	1572
Covariates	Yes	Yes	Yes

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7. Estimates of conditional average treatment effects via OLS interaction terms with anti-immigration beliefs

1.A.8.4 Conditional Average Treatment Effects by Far-Right Voting

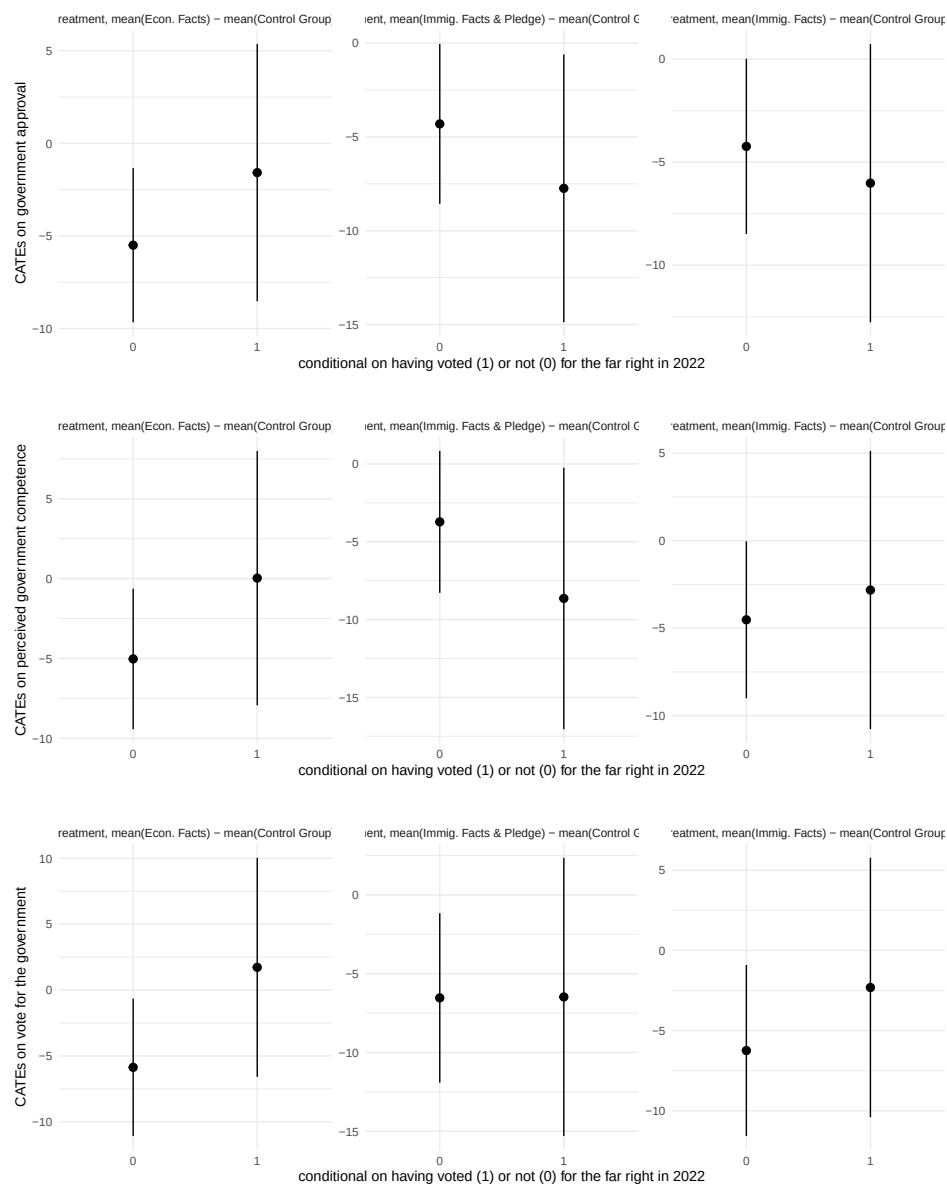


Figure 1.15. Conditional average treatment effects on attitudes to the government w.r.t. whether the respondent voted for a far-right party

	Approval	Competence	Vote Intention
Immig. facts	-4.2† (2.2)	-4.5* (2.3)	-6.2* (2.7)
Immig. facts & pledge	-4.3* (2.2)	-3.7 (2.3)	-6.5* (2.7)
Econ. facts	-5.5** (2.1)	-5.0* (2.2)	-5.9* (2.7)
Far-right voter	34.6*** (3.0)	32.7*** (3.3)	45.3*** (3.7)
Immig. facts x Far-right voter	-1.8 (4.1)	1.7 (4.7)	3.9 (4.9)
Immig. facts & pledge x Far-right voter	-3.4 (4.3)	-4.9 (4.9)	0.1 (5.3)
Econ. facts x Far-right voter	3.9 (4.1)	5.1 (4.6)	7.6 (5.0)
Num.Obs.	1649	1658	1604
Covariates	Yes	Yes	Yes

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.8. Estimates of conditional average treatment effects via OLS interaction terms with far-right voting

1.A.8.5 Most important moderators

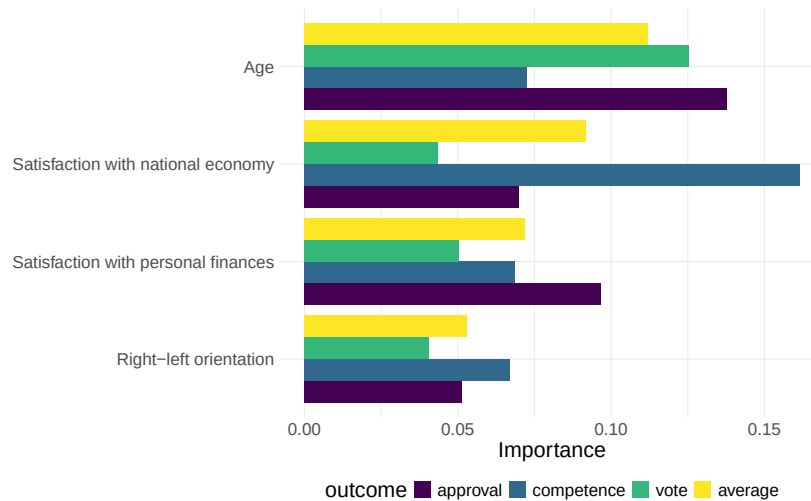
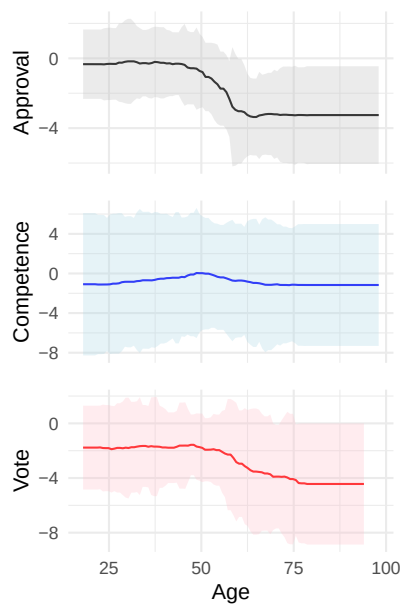


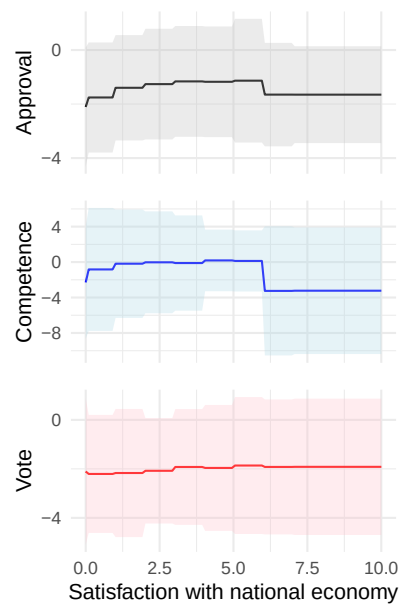
Figure 1.16. Four most important contributors to heterogeneity in treatment effects across outcomes

Although their contribution varies from outcome to outcome, overall, the causal forests identified age, left-right orientation and satisfaction with both the national economy and personal finances as the most important moderators. Figure 1.17 plots the treatment effects conditional on these key variables for each outcome while holding the other covariates at their median values.

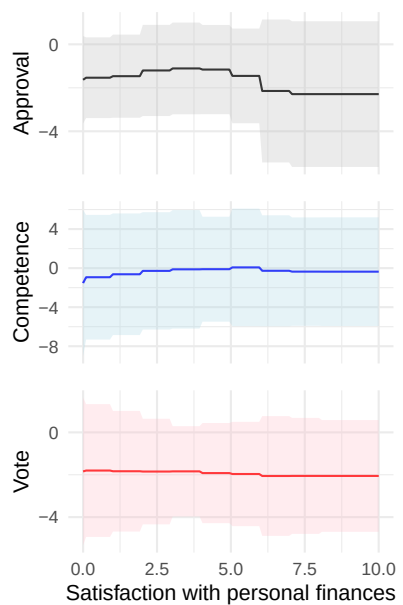
Looking at the point estimates in Figure 1.17, one might suppose that the effects of the immigration treatment are stronger, i.e. more negative, for older and more right-ring respondents, and those who were more satisfied with the national economy or personal finances. However, as the confidence intervals overlap for basically all of these effects, this evidence remains inconclusive. To some degree, this might be related to statistical power, i.e. not having enough observations at every point of age or the Likert scale used for the other questions. However, even the binary variables (providing maximum power when two subgroups are roughly equal) tested with regressions and with causal forest showed no evidence of moderating heterogeneous treatment effects.



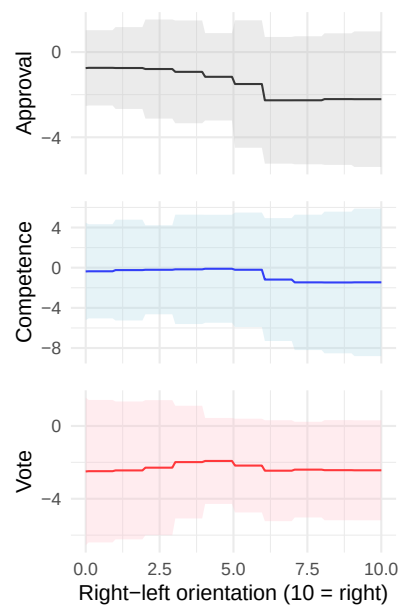
(a)



(b)



(c)



(d)

Figure 1.17. Conditional Average Treatment Effects

1.A.8.6 Respondents Who Reported Trusting the Treatment

After the treatment and the main outcome questions, the treated respondents were asked whether they trusted the statistical information provided by us a few questions ago. Among those who saw the economy vignette, about 43% reported trusting it, while about 59% of those who saw the immigration vignette said they trusted it. Although this question was asked after the treatment and only in the treated groups, in this section, I present suggestive evidence by restricting the treated groups to the subsamples of those who indicated trusting or distrusting the treatment information.

Under the assumption that conditioning on the post-treatment trust in provided information does not distort the estimates, this evidence indicates that the effect was notably stronger among those who trusted the treatment than among those who did not. This intuitively aligns with the logic of treatment working only on some respondents. However, this evidence remains suggestive as it is impossible to causally identify the treatment effects by conditioning on a post-treatment variable.

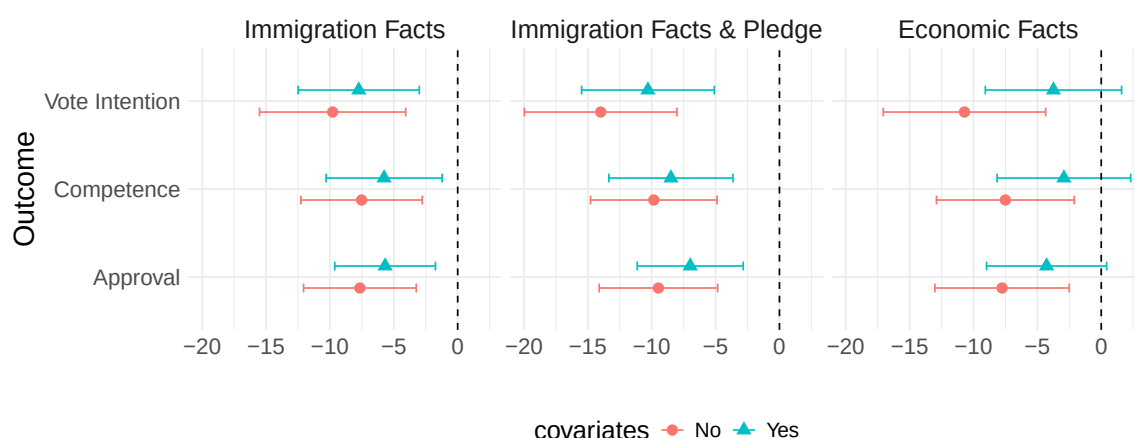


Figure 1.18. Effect estimates limiting the treated to those who reported trusting the treatment

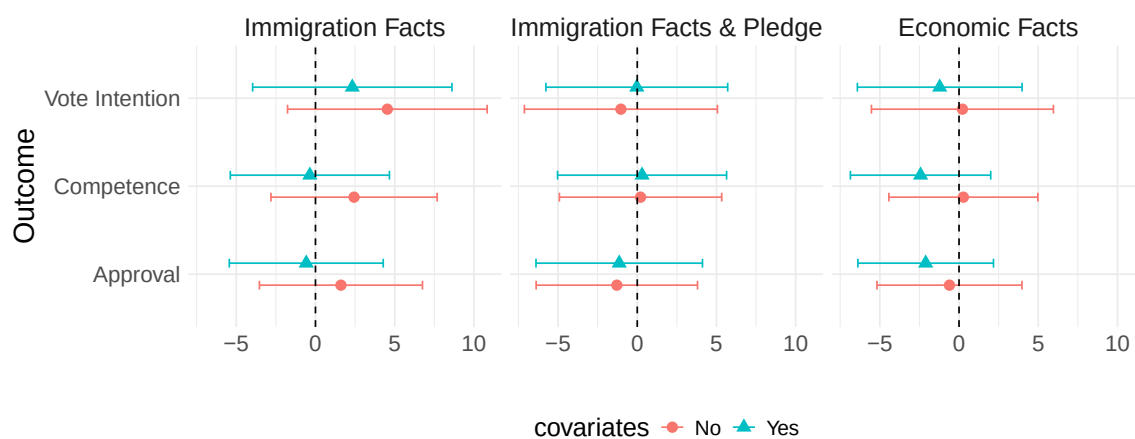


Figure 1.19. Effect estimates limiting the treated to those who reported no trust in the treatment

Chapter 2

Exploiting Divisions: Political Gains from State-Sponsored Homophobia¹

Do strategies of state-sponsored homophobia translate into electoral gains? While a growing body of literature documents the increasing politicization of LGBTQ- and gender-related issues by illiberal elites, little is known about the electoral effects of these strategies. I address this important question by studying whether anti-LGBTQ mobilization pays off electorally for the initiating party. Empirically, I study the adoption of anti-LGBTQ resolutions in many Polish municipalities prior to the 2019 parliamentary election. Using a synthetic difference-in-differences design, I find that these resolutions significantly depressed turnout in affected municipalities, with opposition parties showing less mobilization capacity. By contrast, turnout for the incumbent Law and Justice Party increased substantially. Overall, this study's findings are relevant for understanding the electoral consequences of both elite-led mobilization against stigmatized and discriminated groups, and policies of subnational democratic backsliding.

¹This chapter is based on a paper co-authored with Violeta I. Haas, Tarik Abou-Chadi, Heike Klüver and Lukas Stoetzer, with [individual-level data analyses](#) added by me for the dissertation. I am grateful to my amazing co-authors, as well as to Catherine E. De Vries, Tom Pepinsky, Ivan Aleksandrov, Sara Luxembourg and the participants of the 2023 European Workshop in Empirical Political Science (EuroWEPS), 2023 Elections, Parties and Public Opinion (EPOP) conference, 2023 Italian Political Science Association (SISP) conference, 2024 Southern Political Science Association (SPSA) conference, 2024 Midwest Political Science Association (MPSA) conference and seminars at Bocconi University for their helpful suggestions.

2.1 Introduction

“Our [party] alliance would never be pro-LGBT, because family is sacred to us. We will bury those pro-LGBT in the ballot box”

Recep Tayyip Erdoğan

Speech at an election rally on May 7, 2023 (Reuters, [2023](#)).

This quote by Turkey’s president, Recep Tayyip Erdoğan, exemplifies a broader trend: autocratic leaders and radical right parties have intensified their campaigns against LGBTQ communities. Across the globe, illiberal actors – mostly from the right – have used political homophobia in rhetoric and policy in remarkably congruent ways (Ayoub & Stoeckl, [2024](#); Bosia & Weiss, [2013](#)). Repressive anti-gender and anti-LGBTQ policies have become common practice in many countries, with laws such as “sexual propaganda” statutes enacted in Russia and Hungary, and numerous US states restricting gender-affirming care. In Western Europe and Latin America, radical right parties and religious groups increasingly focus on an anti-LGBTQ agenda (Abou-Chadi et al., [2021](#); Corrales, [2019](#); Magni & Reynolds, [2023](#); Payne & Santos, [2020](#); A. E. Smith & Boas, [2024](#)). While scholars have engaged with the discursive and institutional meaning of state-sponsored homophobia, we know much less about its effects on political behavior.

A growing literature on LGBTQ politics has studied how progressive change such as marriage equality legislation (Abou-Chadi & Wagner, [2019](#); Flores & Barclay, [2016](#)) or pride events (Ayoub & Kollman, [2021](#)) can affect attitudes toward the LGBTQ population. While we have overall seen a remarkable positive shift in attitudes towards homosexuality, a crucial finding of this literature is that elite-led progressive change will not necessarily lead to more progressive and inclusive attitudes but also potentially to backlash (Ayoub, [2016](#)). In addition, stated, progressive LGBTQ attitudes might be instrumental rather than genuine (Turnbull-Dugarte & Ortega, [2024](#)).

Studies on the electoral effects of specifically anti-LGBTQ measures have almost exclusively focused on direct democracy initiatives in the US context (Camp, [2008](#); Campbell

& Monson, 2008; Donovan et al., 2008; Garretson, 2014; Hillygus & Shields, 2005; Lewis, 2005; D. A. Smith et al., 2006). These studies reveal the significant impact of ballot measures seeking to prohibit marriage equality on turnout and vote choice. For instance, Campbell and Monson (2008) demonstrate that anti-marriage equality ballot initiatives in 2004 spurred evangelical Christians to mobilize in support of George W. Bush while diminishing turnout among secular individuals.

In this research note, we argue and empirically demonstrate that adopting measures of state-sponsored homophobia can increase support for authoritarian government parties. We expect that such measures can mobilize supporters of the governing party while at the same time decreasing turnout for the opposition. Ayoub (2020) and Page et al. (2022) show that more inclusive LGBTQ rights are associated with higher levels of political participation by individuals who have more tolerant positions on sexuality. This is linked to perceptions of political efficacy.

Hence, we argue that political homophobia is a powerful tool for authoritarian leaders to affect their fortunes at the polls. Introducing anti-LGBTQ measures can help to suppress turnout among their opponents, while mobilizing their own supporters. This, in turn, should increase their electoral support overall. The effects of such measures go beyond standard ideas of issue-voting. Political homophobia cannot be reduced to a programmatic offer that creates varying party support based on attitudes toward these policies. As Ayoub (2020) demonstrate, political homophobia goes deeper and affects perception of political efficacy itself. Authoritarian parties can thus use these tools to tilt the electoral playing field in their own favor.

We use one of the most prominent cases of contemporary state-sponsored homophobia to study its electoral effects: the establishment of so-called “LGBT-free zones” in hundreds of municipalities in Poland (Rafałowski, 2023). Beginning in March 2019, seven months prior to the parliamentary election, local authorities in parts of Poland started adopting resolutions that declared them disapproving of and free from what they described as “LGBT ideology.” Following our argument, we expect relative decreases in

opposition party turnout and higher mobilization for the governing Law and Justice party (Prawo i Sprawiedliwość, PiS) in “LGBT-free zones.”

The Polish case is ideal to study the electoral effect of political homophobia for several reasons. First, the rare subnational variation in anti-LGBTQ policies mitigates concerns about unobserved confounders, allowing comparisons between regions that are similar to each other within the same country. Second, the uniform treatment of the resolutions avoids pitfalls related to lumping together very different anti-LGBTQ policies. Third, PiS’s explicit association with the promotion and implementation of these resolutions allows for clear responsibility attribution in vote choices.

Our empirical strategy relies on a synthetic difference-in-differences (sDiD) design (Arkhangelsky et al., 2021). Since we know that municipalities did not declare anti-LGBTQ resolutions at random and trends could influence adaptation, we cannot simply compare these municipalities to others in a classic treatment-control framework, or even use standard difference-in-difference approaches. We instead use a sDiD design that allows us to compare municipalities with anti-LGBTQ measures to synthetic ones that had similar pretreatment voting outcome trajectories but have not been treated. In line with our expectations, we study three outcome variables: turnout, opposition party turnout (opposition party votes divided by total eligible votes) and government party turnout. We also present evidence on how these dynamics translate into vote shares.

We find significantly lower turnout in municipalities that passed anti-LGBTQ resolutions compared to the synthetic control group. This is due to the fact that opposition party turnout is significantly lower in these municipalities. Opposition parties were significantly less able to mobilize here. In contrast, party turnout for the governing PiS is significantly higher. These findings hold when limiting our sample to areas in a 50-kilometer (km) radius of treatment-control boundaries, strongly suggesting that we are not capturing broader geographic differences between East and West Poland.

With these findings, we make an important contribution to research on state-sponsored homophobia and LGBTQ politics more generally. We provide the first study

to causally identify an effect of state-led anti-LGBTQ policies on electoral outcomes. We thus contribute to a growing body of work that has demonstrated how sexuality policies shape political behavior. It is especially remarkable that we find behavioral effects of “soft laws” that in principle only have symbolic character.

Our article also contributes to a better understanding of the mechanisms behind democratic backsliding. Anti-gender and anti-LGBTQ policies have become an important part of the toolkit of wannabe autocrats. We can show how these policies, indeed, help authoritarians electorally. As such, this study also connects to the literature on subnational authoritarianism and subnational democratic backsliding (Gibson, 2005; Grumbach, 2023; O’Dwyer & Stenberg, 2022) by demonstrating the effects of local-level institutionalization of illiberal agendas and policies. Political homophobia should thus be seen more clearly as part of the authoritarian tool kit to tile the electoral playing field.

2.2 Background

2.2.1 Political Homophobia in Poland

Political homophobia in Poland, as in many other countries, is not a new phenomenon. During the communist era, the state and society systematically repressed gay individuals, enforcing elevated ages of consent, engaging in secret police persecution, and perpetuating pervasive social stigma (O’Dwyer, 2012). Even before the communist regime ended, the Catholic Church had already emerged as a powerful authority reinforcing homophobia in Poland, framing homosexuality as a moral threat to traditional family values (Ayoub, 2016). However, the topic was largely sidelined from the political debate during the 1990s (O’Dwyer, 2012), a period dominated by the challenges of democratic and economic transition.

With the EU accession process in the early 2000s, LGBTQ issues in Poland became politicized (Ayoub, 2016). While LGBTQ activists gained public visibility and increasingly claimed rights and protections for the community, the 2000s also witnessed a back-

lash and a rise in political homophobia (O'Dwyer, 2012). The framing of homosexuality as a moral failing and a threat to national identity was readily adopted by anti-LGBTQ forces (Ayoub, 2016). Equality marchers have been repeatedly attacked and injured by countermobilizations, denied police protection and also banned in Warsaw in 2004 and in Poznan and Warsaw in 2005 (Ayoub, 2016)². The PiS-led national government also attempted to ban “homosexual propaganda” in schools in 2007. While no further anti-LGBTQ policies were pursued before PiS returned to power, little progress in LGBTQ rights was made during the first two Tusk governments (2007-2014).

After its return to office, PiS repeatedly engaged in anti-LGBTQ rhetoric, including statements from high-profile politicians starting from 2018 (Graff & Korolczuk, 2022). In 2018, the government also pressured 211 schools to withdraw from a “Rainbow Friday” event aimed at supporting LGBTQ students’ wellbeing (ILGA-Europe, 2019).

The anti-LGBTQ resolutions starting in 2019 can be understood as escalating these earlier strategies. At the same time, civil society responses, including the efforts of grassroots activists and legal challenges at the European level, underscore the contested nature of political homophobia in Poland.

2.2.2 Anti-LGBTQ Resolutions

Beginning in March 2019, seven months prior to the parliamentary election, local authorities in Poland started adopting resolutions that declared them disapproving of and free from what they describe as “LGBT ideology.” Internationally, these regions became known as “LGBT-free zones,” a term popularized by both right-wing actors, like the weekly magazine *Gazeta Polska* (Bucholc, 2022), and activists who strategically positioned city signs in affected regions to highlight the social and political ramifications of these policies (Stenberg & O'Dwyer, 2023). The associated documents were widely perceived as a direct response by PiS to the earlier signing of the “LGBT+ Declaration for Warsaw” by the city’s mayor and later presidential candidate of the Civic Platform, Rafał Trzaskowski (Rafałowski, 2023). Although the resolutions were symbolic and legally

²Another ban was attempted in Lublin in 2018 and overturned by a court.

non-binding, defended by provincial legislatures as neither acts of local law nor public administration and as imposing no obligations or conferring any rights (Świętokrzyskie Voivodeship Assembly, 2019), they nonetheless raised serious human rights concerns regarding the safety and well-being of Poland's LGBTQ population. By encouraging local governments to refrain from promoting tolerance, supporting LGBTQ rights organizations, or implementing anti-discrimination education, the resolutions intensified minority stress with tangible harm. Most dramatically, recent research indicates that anti-LGBTQ policies in Poland increased suicide attempts by 16 percent (Meyerhoefer et al., 2025).

These resolutions were passed at three different administrative levels - municipalities (gminas), counties (powiats) and provinces (voivodeships) - and with a few customized exceptions were generally presented in two different variants - "Charter of the Rights of the Family" and "Resolution against LGBT ideology." While resolutions were adopted on local levels, their striking resemblance and synchronized adoption in PiS-controlled areas have been interpreted as evidence of a "nationwide, coordinated anti-LGBTQ campaign" (Bogatyrev & Bogusz, 2025, p. 2).

Anti-LGBTQ resolutions were typically initiated by two groups of documents. The first and most prevalent group of resolutions consists of those opposing and proclaiming freedom from "LGBT ideology." They contain statements denouncing "political correctness," condemning efforts to educate about non-heterosexual relationships, and include warnings against radicals seeking "cultural revolution" (Bucholc, 2022). The second variant, titled "Local Government Charter of The Rights of The Family," features more nuanced appeals for preserving "traditional" family models and shielding children from the influence of "various ideologies" that challenge the established social order, clearly alluding to LGBTQ issues (Rafałowski, 2023). The Charter was drafted and lobbied for by the Ordo Iuris Institute for Legal Culture, an ultra-conservative think tank and advocacy group that thrived during the PiS government and played a significant role in shaping policies against abortion, sex education, and LGBTQ rights (Hennig, 2023). Notwithstanding slight variations in wording and titles across administrative units, both types of resolutions exhibit clear aversion towards non-heteronormative relationships and

family models. In addition, both variants can be understood as soft laws, i.e., agreements, principles, and declarations that are of symbolic significance rather than legal force, do not entail any binding obligations and are thus unable to be enforced in court.

Shortly after the parliamentary election 2019 and the release of an interactive map by Polish LGBTQ activists documenting the spread of anti-LGBTQ resolutions, international outcry over these policies grew rapidly. In December 2019, the European Parliament (EP) passed a resolution, condemning the anti-LGBTQ resolutions and urging action to safeguard the rights of LGBTQ people in Poland. The EP's resolution also called on EU authorities to use tools such as infringement procedures and budgetary procedures "to ensure the full and proper application of Treaty principles and values" (Parliament, 2019). International pressure continued to mount, when numerous cities across Western Europe threatened to terminate their partnerships with their concerned sister cities in Poland (Rafałowski, 2023).

Despite these efforts, local councils persisted in passing resolutions in the month following the 2019 parliamentary election. By mid 2020, even larger parts of the Polish population resided in areas where anti-LGBTQ proposals had been approved. As such, the adoption of anti-LGBTQ resolutions follows a wave-like pattern, wherein a first round of municipalities (N=741) was covered by resolutions enacted before the 2019 parliamentary election, and a second, smaller round of municipalities (N=164) was affected by resolutions passed after the 2019 parliamentary election. Figure 2.1 shows the geographic coverage of anti-LGBTQ resolutions by wave. The overlap of these areas with PiS strongholds in the 2019 election is remarkable (see Figure 2.2).

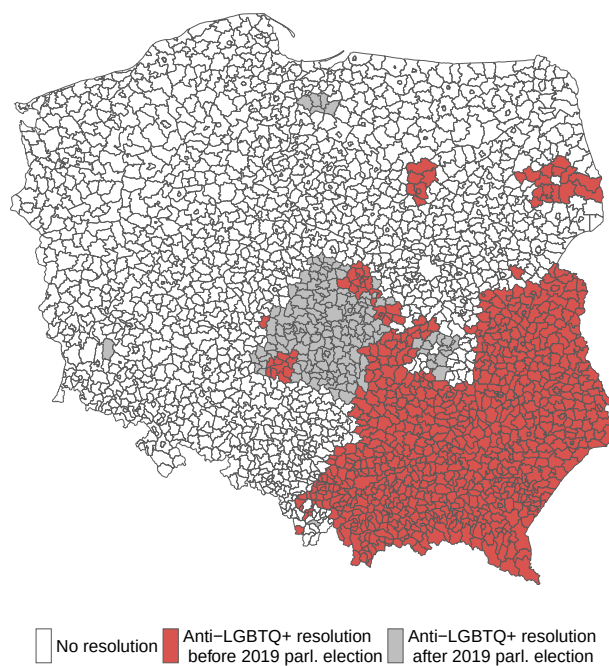


Figure 2.1. Geographic distribution of municipalities with anti-LGBTQ resolutions by wave.

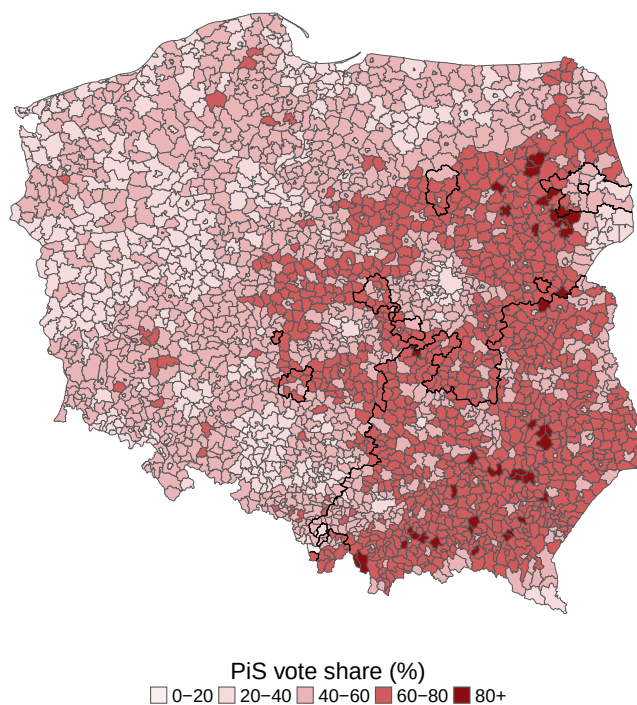


Figure 2.2. Geographic distribution of PiS vote share in the 2019 parliamentary election. Solid lines indicate the boundaries of anti-LGBTQ resolution coverage.

In July 2020, the EU followed through on their warnings, denying funding to six Polish towns that had adopted anti-LGBTQ resolutions (Wanat, 2020). This constituted a major threat for Poland, which, at the time, was the largest net beneficiary of EU funds and relied heavily on financial support. In response to the heightened risk of funding rejections and potential EU infringement proceedings, nearly all local anti-LGBTQ resolutions were ultimately rescinded before the 2023 parliamentary election (Bogatyrev & Bogusz, 2025).

2.2.3 LGBTQ issues in the 2019 Elections

By the time of the 2019 parliamentary election, a quarter of the Polish population resided in areas where either of the two anti-LGBTQ proposals had been approved. Figure 2.3 shows the geographic coverage of anti-LGBTQ resolutions prior to the 2019 parliamentary election. Adoptions were primarily concentrated in conservative, rural regions to the southeast of the country.

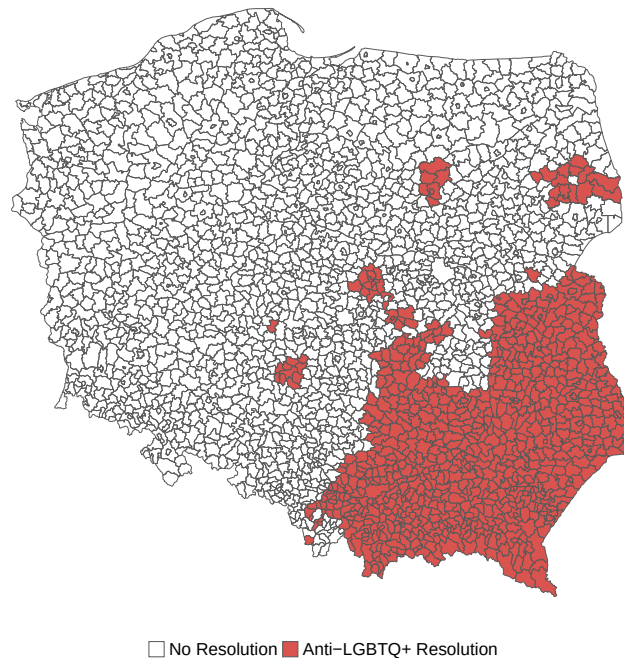


Figure 2.3. Geographic distribution of municipalities with anti-LGBTQ resolutions before the 2019 parliamentary election.

The parliamentary election held in October 2019 was fiercely contested, with LGBTQ rights emerging as a highly salient issue. The ruling PiS party, in particular, made its devotion to “traditional” family models and opposition to LGBTQ rights one of the main themes of its campaign. For instance, in April 2019, PiS party chairman Jarosław Kaczyński labeled the LGBTQ rights movement a “threat” to Polish identity and the state (Associated Press, 2019). PiS also portrayed itself as the defender of “traditional” family values against this perceived threat (Graff & Korolczuk, 2022). Illustrating this stance, the then-president of the Polish Senate, Stanisław Karczewski, shared an image on Twitter depicting PiS as a protective umbrella for the heteronormative Polish family amidst a storm of rainbow-colored rain (see Appendix Section 2.A.1). In contrast, the centrist-liberal Civic Coalition (KO), as the main opposition force, expressed support for LGBTQ rights, condemning the enactment of anti-LGBTQ resolutions. Meanwhile, the left-wing alliance (Lewica) also advocated for LGBTQ rights in their campaigns. The centrist-conservative Polish Coalition (KP), on the other hand, being PiS’s main competitor in rural constituencies, has largely avoided LGBTQ issues in their campaign, while the far-right challenger Confederation (Konfederacja) echoed the anti-LGBTQ platform of the ruling PiS party (Gwiazda, 2023).³

The election resulted in a landslide victory for PiS, which captured 43.6% of the popular vote, the most substantial mandate attained by any party since Poland’s transition to democracy. At the same time, voter turnout surged, hitting a three-decade high of 61.7%.

2.3 Data and Method

2.3.1 Data on Anti-LGBTQ Resolutions

To construct the treatment variables based on the adoption of anti-LGBTQ resolutions in particular parts of Poland, we rely on the “Atlas of Hate” (Pająk & Gawron, 2020)

³Further information on party positions towards LGBTQ rights in 2019 can be found in Appendix Section 2.A.2.

as the data source. The Atlas of Hate is an interactive online map of Poland plus a publicly available underlying dataset, which has been maintained by Polish human rights activists since 2019. Due to the way the data gets updated in case of repeals (by erasing the previous status), we use several versions of the same dataset downloaded at different times since 2020 to construct the full data.

This dataset provides extensive details about the status (adopted/not adopted/retracted), administrative unit (municipality/county/province), type and date of each anti-LGBTQ resolution passed in Poland since 2019. Virtually every case also includes links to the official sources of the resolution text, meeting agenda, minutes, video record and voting results on the resolution as well as press information about the decision. This allows for verification of each case through public sources. Our own cross-references of randomly sampled entries have validated the accuracy of the data.

The reliability of the dataset is corroborated by the reputation of the “Atlas of Hate”. In 2020, the project was nominated for the EU’s most prestigious human rights award – the Sakharov Prize (European Parliament, 2020). Even within Poland, the credibility of the “Atlas of Hate” was supported by court decisions against local authorities who sued to get removed from the map (Ptak, 2022). Notably, the “Atlas of Hate” data have been previously used for academic research (Bogatyrev & Bogusz, 2025; Meyerhoefer et al., 2025; Rafałowski, 2023; Stenberg & O’Dwyer, 2023).

2.3.2 Treatment

To identify regions with anti-LGBTQ resolutions in place, we rely on the “Atlas of Hate,” an interactive map of Poland developed by activists (Pająk & Gawron, 2020).⁴ Defining the treatment is not straightforward. Resolutions were adopted across various levels of government using two different document types. The “Atlas of Hate” provides detailed information about the status (adopted/not adopted/retracted), subnational unit (municipality/county/province), type (“Resolution against LGBT ideology”/“Charter of the Rights of the Family”), and date of each instance. For our main analysis, we rely on data

⁴For details on this data source and its quality, see Appendix 2.3.1.

at the municipality level, defining a municipality as treated if either its local council or a superior subnational unit passed any anti-LGBTQ resolution.

The adoption of resolutions through two distinct documents and across varying (sometimes multiple) levels of government creates conditions for potential heterogeneity in treatment effects based on signaling strength and mobilization dynamics. On the one hand, the “Resolution Against LGBT Ideology” may constitute a stronger treatment, given its explicit targeting of LGBTQ communities. On the other hand, the adoption of resolutions at lower government levels could reflect a process of grassroots mobilization. By contrast, in cases where municipalities were affected by adoptions on higher-level entities (i.e., county or province level), mobilization dynamics may have been predominantly top-down. This distinction may plausibly have implications for the salience of these resolutions across treated municipalities. Specifically, we might expect voters in municipalities directly adopting resolutions to be better informed about the initiatives compared to those in municipalities subjected to top-down implementation. Similarly, municipalities affected by both types of resolutions or across multiple administrative levels may have received a stronger signal. We address these considerations in the robustness section, where we evaluate variation in outcomes by resolution type and administrative level separately, as well as by cumulative adoption patterns.

Additionally, in a small subset of cases, anti-LGBTQ resolutions were proposed but not adopted. In these municipalities, resolutions were discussed (debated but not brought to a vote), rejected (voted on but ultimately not passed), or dismissed (placed on the agenda but neither debated nor voted on). A potential concern is that including these cases in the control group may contaminate it, contributing to any observed differences between the treatment and control groups. We address this issue in the robustness section.

2.3.3 Outcomes

We collect municipality-level data for parliamentary elections from the National Electoral Commission. To study pre-treatment trends in voting behavior, we use data on all parliamentary elections going back to 2001, when PiS was founded. We focus on three primary

outcomes. First, we consider overall turnout, defined as the number of ballots received divided by the number of eligible voters. Second, we present the combined vote shares for opposition party lists relative to the vote-eligible population (i.e., opposition turnout) to facilitate interpretations of (de)mobilization patterns. Finally, we also calculate government turnout (i.e., PiS vote share relative to vote-eligible population). While traditional measures of vote shares ignore turnout, offering insights solely into the relative proportions of votes, analyzing opposition/government-specific turnouts allows us to directly test the mobilizing and demobilizing effects of anti-LGBTQ resolutions. We show further evidence on vote shares in Appendix section 2.A.6.⁵

2.3.4 Sample

The full sample includes all 2,477 municipalities in Poland, observed across six parliamentary elections held between 2001 and 2019 ($N = 14,862$).⁶ Among these, 741 municipalities were affected by anti-LGBTQ resolutions during the 2019 parliamentary election, while the remaining 1,736 municipalities were untreated. We exclude the untreated city of Warsaw due to its unique status as the capital, with a population of 1.8 million, making it hardly comparable to any single Polish municipality. Furthermore, votes cast by Polish citizens residing abroad and by personnel stationed on maritime vessels are omitted. As a result, the sample of analysis consists of 14,856 municipality-year observations.

2.3.5 Identification Strategy

Evaluating the causal effect of introducing anti-LGBTQ resolutions on our outcomes of interest comes with a key challenge. The observed spatial patterns of municipalities that adopted anti-LGBTQ resolutions are hardly exogenous or random. Treated regions are

⁵For party alliances where individual parties ran on separate lists or were nonexistent in elections prior to 2019, we defined pre-treatment outcomes as the vote share of the leading party. In addition, while theoretically appealing, including Confederation as the fifth party alliance running in 2019 is empirically unfeasible due to insufficient pre-treatment observation periods in outcomes.

⁶Since 2001, various regional reforms have altered the boundaries of municipalities, primarily through mergers. To ensure comparability across time, we reconstructed the 2,477 municipalities as of 2019 for the parliamentary election of 2001, 2005, 2010, and 2015 by retrospectively merging municipalities. In the event of a split, we retrospectively divided all absolute vote indicators proportionally to the vote-eligible populations of each part.

heavily clustered in the more conservative PiS strongholds within the southeast of Poland (see Figure 2.3). These regions differ considerably from the rest of the country in ways potentially threatening assumptions of traditional causal inference designs, like difference-in-differences (DiD).

Standard DiD estimation compares the changes in vote shares and turnout before and after the introduction of resolutions in municipalities that implemented them (treatment group) versus those that did not (control group), and then calculates the difference in changes between the two. The identification assumption of DiD designs requires parallel trends, which implies that in the absence of anti-LGBTQ resolutions, both groups of municipalities would have experienced similar trends in voting. In practice, verifying this assumption is impossible, and it may be violated when pre-existing differences between treatment and control municipalities affect post-exposure outcomes. This is likely to be the case with our geographically clustered treatment. Previous evidence shows that anti-LGBTQ resolutions were more likely to be proposed and adopted by PiS-controlled local governments (Rafałowski, 2023; Stenberg & O'Dwyer, 2023). Indeed, raw data seldom supports parallel trends between treatment and control groups, and our case is no exception. The red and dotted grey lines in Figure 2.4 show the observed (unweighted) pre-treatment trends of the three primary outcome variables - overall turnout and government/opposition turnout - for all parliamentary elections between 2001 and 2019, by treatment group. Neither for overall turnout nor opposition and government turnout do pre-treatment trends move in parallel between observed groups. Both levels and trajectories in the pre-treatment period show clear discrepancies. In such scenario, the observed control group may not serve as an accurate counterfactual for the treated group.

We address this concern in two ways. First, we use synthetic difference-in-differences (sDiD), a generalized version of the DiD and synthetic control methods that weakens the reliance on parallel trend type assumptions (Arkhangelsky et al., 2021). The sDiD approach constructs a synthetic control group that closely resembles the treatment group. Specifically, synthetic controls are generated using a weighted mixture of observed control units that are similar to the treatment group in terms of pre-intervention outcome

trends.⁷ Second, we repeat our analysis on a restricted sample of municipalities within a 50 km radius of treatment-control boundaries.⁸ By limiting our sample to communities around the treatment-control boundary, we ensure that the observed effects are primarily attributable to anti-LGBTQ resolutions rather than simultaneous but unrelated events happening in the southeast of Poland. Furthermore, this approach also addresses potential confounding factors related to the region’s historical and path-dependent political legacies, which continue to influence varying levels of alignment with liberal values today (Charnysh, 2024; Grosfeld & Zhuravskaya, 2015, e.g.).

As described in Appendix Section ??, resolutions continued to be passed after the 2019 parliamentary election, but were largely rescinded by the time of the 2023 parliamentary election (Bogatyrev & Bogusz, 2025). Our approach for generating synthetic control weights draws on all municipalities that had not yet implemented anti-LGBTQ resolutions by the 2019 parliamentary election. Including control units that enacted resolutions post-2019 parliamentary election presents a hard empirical test, as it requires effects to hold even when compared to municipalities that subsequently selected into treatment.⁹ Moreover, to address the fluctuations in treatment caused by the subsequent withdrawal of resolutions, we refrain from analyzing the effects on the 2023 parliamentary election. In doing so, we recognize that outcomes in formerly treated municipalities are likely shaped by the subsequent rejection of EU funding and the repeal of resolutions rather than the initial adoptions at this point in time. The blue line in Figure 2.4 displays the synthetic trends over time, calculated using the weighted average of control municipalities. The graphical evidence illustrates how the sDiD approach mitigates concerns about parallel trends and aligns the time trends of unexposed control units with those of treated municipalities through re-weighting. In what follows, we report on the average treatment effect on the treated (ATT) that is, the average effect of introducing an anti-LGBTQ resolution among receiving municipalities. Following Arkhangelsky et al. (2021), we use jackknife variance estimation to calculate standard errors.¹⁰

⁷Appendix Section 2.A.3 shows the geographic distribution of synthetic control weights by outcome.

⁸Appendix Section 2.A.4 shows a map of municipalities within this 50 km distance. The results remain robust across various bandwidths.

⁹We note that results remain robust when using never-treated units (see Appendix Section 2.A.9.1).

¹⁰Additionally, in Appendix Section 2.A.10.1, we report results using the “placebo method,” instead. The standard errors show only minimal variation across specifications.

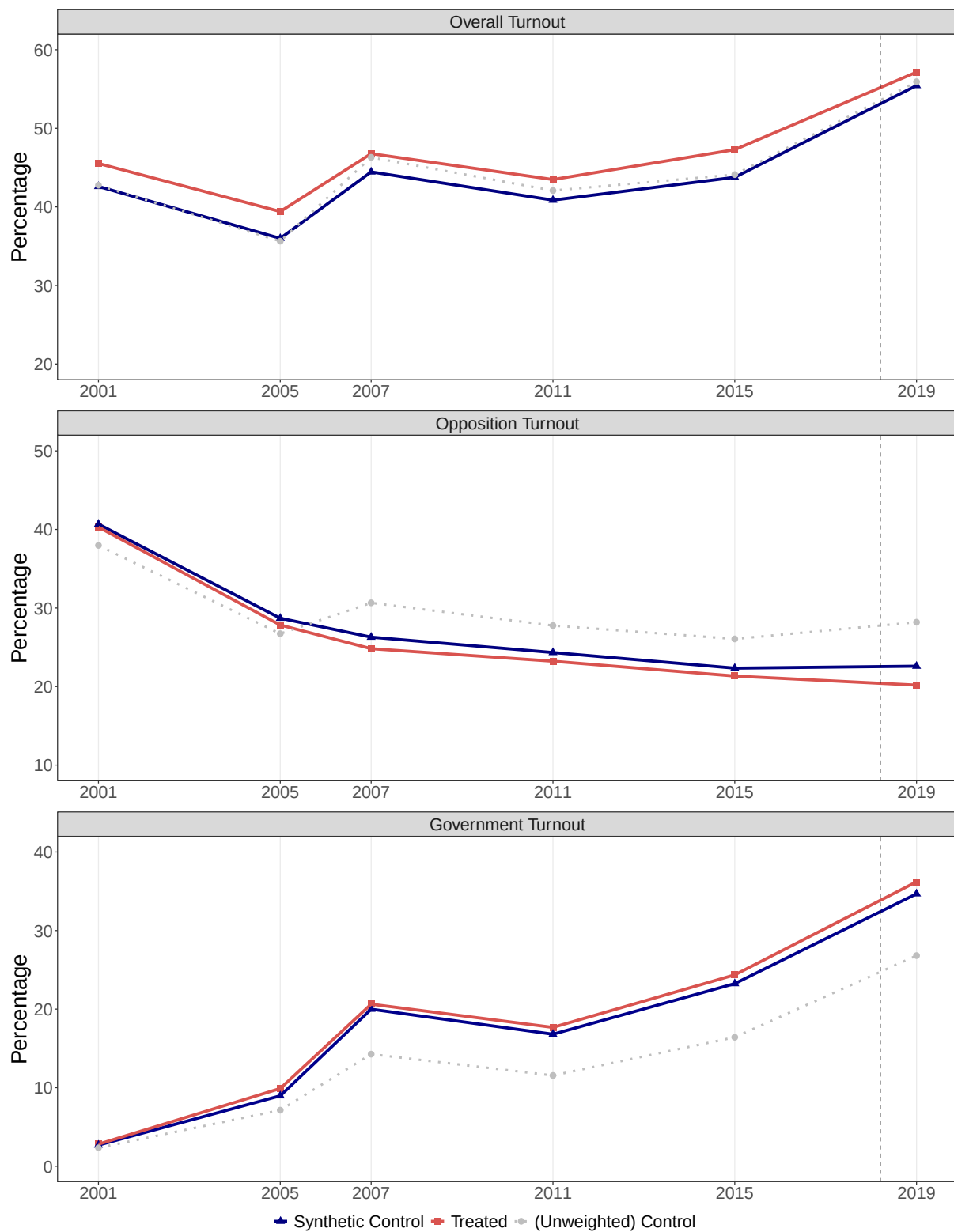


Figure 2.4. Average overall, opposition, and government turnout (%) among municipalities for all parliamentary elections between 2001 and 2019 with the red line indicating treated, the dotted grey line observed (unweighted) control, and the blue line synthetic control municipalities.

2.4 Results

2.4.1 The Electoral Implications of Anti-LGBTQ Resolutions

Figure 2.5 shows the sDiD estimates for overall, opposition, and government turnout for the 2019 parliamentary election. The estimates from both the full sample and the restricted sample within a 50 km radius of treatment-control boundaries generally align in magnitude and direction.

Starting with overall turnout, we find that municipalities adopting anti-LGBTQ resolutions experienced a substantial and statistically significant negative effect on participation during the 2019 parliamentary election. Municipalities with anti-LGBTQ resolutions in place registered changes in turnout that were, on average, 1.69 percentage points lower than those in our synthetic control group. Although voter turnout generally surged in the 2019 parliamentary election (see Figure 2.4), this trend was less pronounced in areas that adopted anti-LGBTQ resolutions.

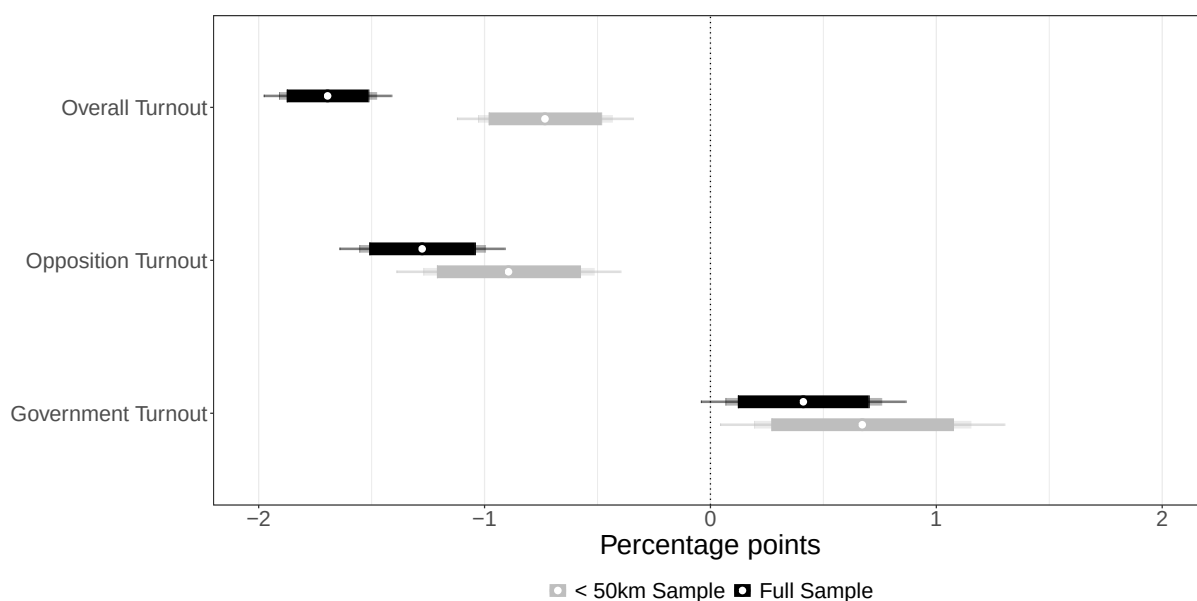


Figure 2.5. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout in the 2019 parliamentary election by sample. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Section 2.A.5.

Moving to our main outcomes of interest, evidence suggests that the observed negative effect on overall turnout within municipalities adopting anti-LGBTQ resolutions is predominantly driven by the relative demobilization of opposition party supporters. Turnout for opposition parties, including KO, KP, Lewica, and Konfederacja, was, on average, 1.28 percentage points lower in these municipalities compared to the synthetic control group. At the same time, the ruling PiS party appears to have secured a mobilization advantage from the adoption of anti-LGBTQ resolutions despite the overall decline in voter turnout observed in affected municipalities. Government turnout in these areas was, on average, 0.41 percentage points higher compared to the synthetic control group, meaning that PiS successfully mobilized a larger share of its potential electorate where such resolutions were enacted. These dynamics translate into overall higher vote shares for PiS. Municipalities with anti-LGBTQ resolutions experienced, on average, a 0.82 percentage point increase in support for PiS (see Appendix Section 2.A.6). This underscores the potential for such measures to significantly alter electoral outcomes, particularly within Poland’s highly polarized political environment.

The meaningful magnitude of these findings is further reinforced when the effects are expressed as percentage reductions relative to pre-treatment turnout levels in the affected municipalities. The estimated 1.69 percentage point decline in overall turnout corresponds to a 3.57% reduction. Similarly, opposition turnout experienced a 6% reduction.

2.4.2 Additional Analysis and Robustness Checks

To test expectations concerning heterogeneity in treatment effects stemming from signaling strength and mobilization dynamics across different (and multiple) resolution types and administrative levels, we perform several additional tests.

We first account for heterogeneity in signaling strength by examining the effects of cumulative exposure, defined as the sum of resolutions covering a municipality or the sum of administrative levels at which a municipality is treated (see Appendix Section 2.A.7). Consistent with theoretical expectations, municipalities exposed at multiple levels or by multiple resolution types reveal the strongest negative effect on turnout. For opposition

and government turnout, estimates also generally align with these patterns. However, the number of observations becomes small for cases involving cumulative exposure, limiting precision.

Second, in Appendix Section 2.A.8, we present results using alternative treatment definitions to capture potential variation in both mobilization dynamics and signaling strength. We begin by examining the impact of defining treatment at distinct levels of adoption, i.e., resolutions passed at the municipality, county, or province level only (see Figure 2.14 in the Appendix for a graphical illustration of this test). While treatment effects exhibit minimal variation across administrative levels, the most substantial effects on both turnout as well as opposition and government turnout are observed at the municipality level. Again, statistical significance diminishes considerably at the municipality level due to the relatively small sample size, with only 52 out of 2,477 municipalities directly adopting anti-LGBTQ resolutions. Overall, these findings provide empirical support for the theorized distinction between grassroots-driven and top-down mobilization processes. Additionally, separate analyses by resolution type reveal comparable effects for both the “Resolution Against LGBT Ideology” and the “Local Government Charter of the Rights of the Family,” suggesting that neither resolution type emerges as the primary driver of the observed outcomes.

Finally, we perform three additional tests to assess the robustness of our findings. First, in Appendix Section 2.A.10.1, we report results from alternative estimators, namely simple Difference-in-Differences (DiD) and Matrix Completion (MC) (Athey et al., 2021). Second, we examine the robustness of our results by excluding the ten largest municipalities with over 300,000 inhabitants, addressing the likely large heterogeneity in election outcomes within these units (see Appendix Section 2.A.10.2). Third, we re-run the analysis with a refined control group, excluding municipalities that discussed, rejected, or dismissed resolutions (see Appendix Section 2.A.10.3). Estimates from alternative estimators for overall, opposition, and government turnout consistently support our findings. Moreover, results remain unchanged when excluding the largest municipalities or municipalities that proposed but did not adopt resolutions.

In summary, our findings paint a worrying picture of how anti-LGBTQ resolutions affect electoral outcomes. The findings suggest that anti-LGBTQ resolutions demobilized opposition voters and benefited the governing radical right PiS.

2.4.3 Individual Level Evidence

To get a better understanding of where the observed relative gains in PiS vote share from anti-LGBTQ resolutions might stem from, we use the 2019 Polish National Election Study (Markowski et al., 2023). Conducted shortly after the 2019 elections, the PNES was a face-to-face survey of 2,003 adult Polish citizens, representing the vote-eligible population of Poland. The survey asked respondents about the vote choices they made in the 2019 and 2015 parliamentary elections, alongside a battery of questions on political attitudes and demographics.

Net voter flow diagrams in Figure 2.6 provide the following descriptive insights into differences in vote-switching dynamics between municipalities affected and unaffected by anti-LGBTQ resolutions. Comparing the two diagrams suggests that in the treatment group, PiS was able to attract more voters from KO than in the control group. Meanwhile, KO was better able to mobilize non-voters in treated than in control municipalities, thus compensating its losses to other parties. Consistently with our sDiD evidence, in places affected by anti-LGBTQ resolutions, the Left experienced less significant growth, gaining less from previous KO-voters and non-voters, who were more likely to vote for the Left in areas unaffected by anti-LGBTQ resolutions.

Additionally, we employ a multinomial logistic model, regressing 2019 individual electoral choices on their previous choices, a binary indicator of treated municipalities of residence and a battery of demographic covariates. Figure 2.7 presents the marginal effects of living under an anti-LGBTQ resolution on the choice made in the 2019 elections, conditional on the 2015 voting behavior.

Overall, the marginal effects are small and most do not reach statistical significance. This is no surprise, given the already modest effects observed at the aggregate level, which

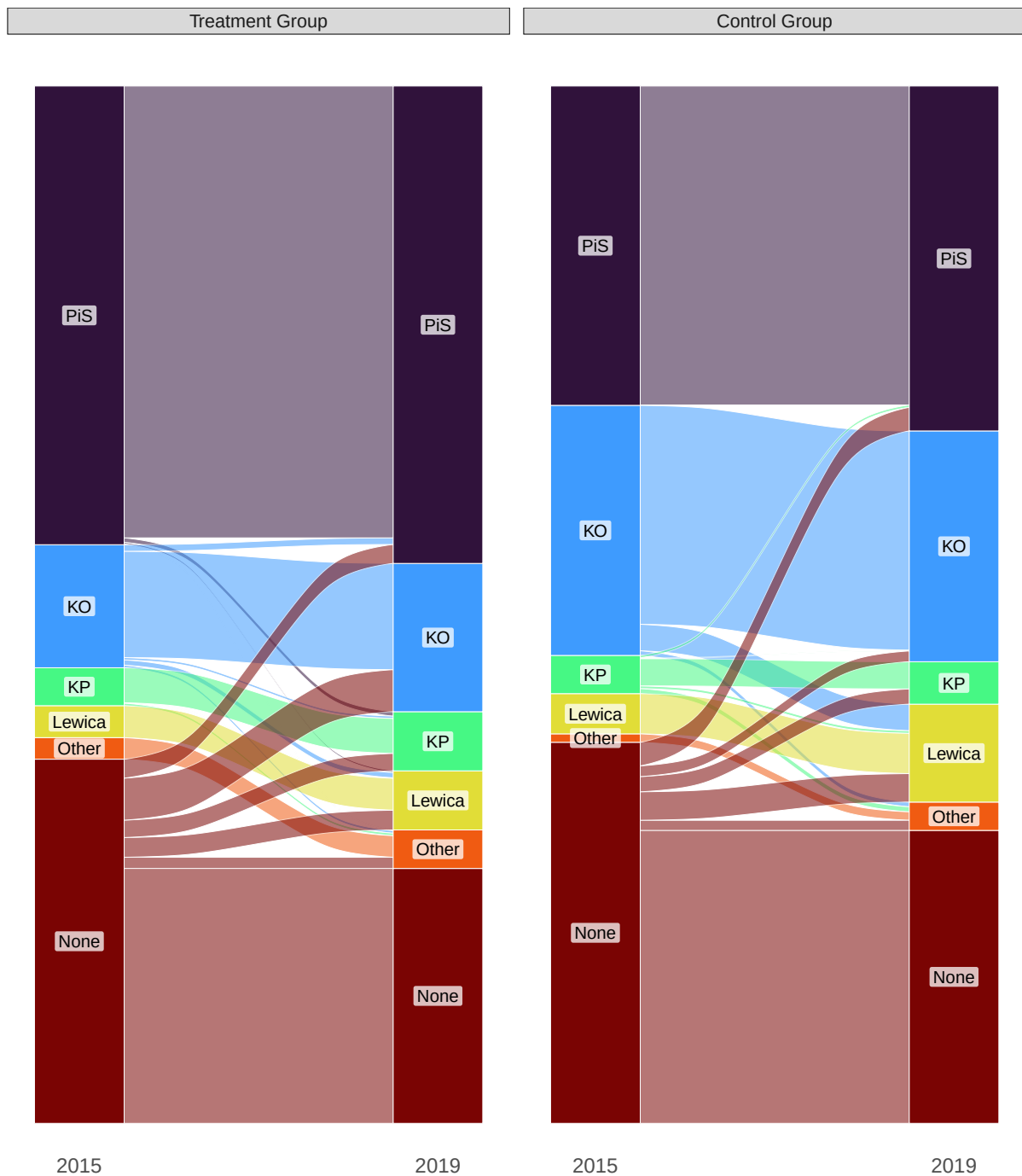


Figure 2.6. Net voter flow diagrams based on the Polish National Election Study 2019. The net flows are obtained by subtracting the number of voters switching from one party to another from the larger number moving in the opposite direction. The 2019 party alliances were reconstructed for 2015 from constituent parties.

makes effects on the individual level difficult to detect amidst sample size constraints posed by the PNES. The confidence intervals of the marginal effects are about 10-15 percentage

points wide, which is too imprecise to detect a likely minor proportion of voters whose behavior was affected by anti-LGBTQ resolutions. Speculating on the point estimates of the individual-level marginal effects suggests that anti-LGBTQ resolutions made non-PiS voters more likely to switch parties (including switching to PiS) and abstainers more likely to keep abstaining (thus, staying demobilized). Meanwhile, the zero effects among previous PiS voters underscore that anti-LGBTQ resolutions did not affect their loyalties, letting PiS keep their constituents and benefit electorally from the effects on other voters.

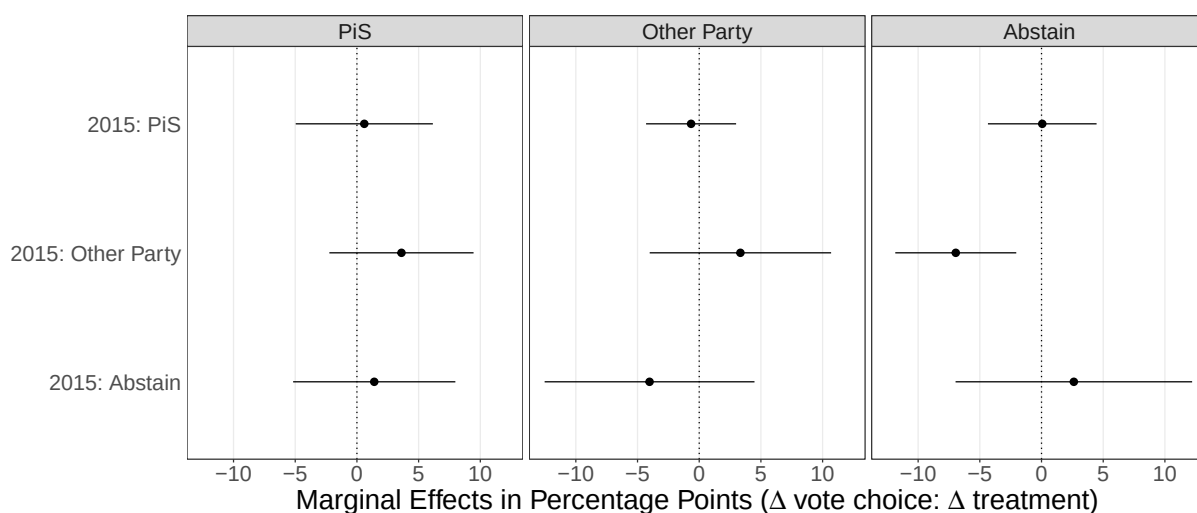


Figure 2.7. Marginal Effects on abstaining, voting for PiS, or voting for any other party in the 2019 parliamentary election, conditional on reported vote choice in the 2015 parliamentary election. Marginal effects show the difference in the probability of switching for respondents residing in municipalities with anti-LGBTQ resolutions in place, compared to those in control municipalities.

2.5 Conclusion

In this study, we provide a first systematic test that causally identifies the electoral consequences of state-sponsored homophobia. Using the case of Poland, we find that anti-LGBTQ resolutions increased support for the governing party, PiS, in the 2019 parliamentary election. PiS was able to mobilize higher levels of turnout among its supporters, while turnout for opposition parties overall decreased. These findings are in line with Ayoub (2020), who demonstrate the important role of sexuality politics for perceived political

efficacy and turnout.

Our findings thus provide an important insight into global anti-gender and anti-LGBTQ strategies. Simply put: these strategies work. We can causally identify an effect of exclusionary anti-LGBTQ measures on mobilizing support for authoritarian governments. Beyond the purely academic assessment, this should be worrying for two reasons. First, if political homophobia is indeed an electorally beneficial strategy - and has remained one despite overall more progressive attitudes toward sexual minorities - then we will likely see more of these strategies. The current wave of anti-LGBTQ measures would then not be a backlash but the start of a new era of state-supported exclusion and oppression. Importantly, domestic contexts may affect how these signals affect political behavior. Further research is needed to causally identify the effects of state-sponsored homophobia in, for example, more liberal societies.

Second, the fact that authoritarians can rile up support through political homophobia is not only bad news for the groups most adversely affected by these policies, but also for democratic support more generally. Autocrats around the globe have found an instrument that seems to work to create additional support. This is even more worrying as we have studied the effects of a “soft law” largely of symbolic nature. Such measures are relatively easily implemented and have become a common strategy among radical right politicians. They complicate oversight and sanctioning from international actors and are more difficult to challenge in court. National and international proponents of liberal democracy, thus, need to find ways to address state-sponsored homophobia, not only in the interest of marginalized minorities but also liberal democracy itself.

That said, the Polish case also provides some reason for optimism. The recent electoral defeat of PiS in the 2023 parliamentary election shows that anti-gender and anti-LGBTQ strategies may not yield benefits in the long-run. These electoral losses, possibly fueled by (international) pressure from LGBTQ activists and, eventually, EU authorities, may thus also suggest one potential avenue for addressing strategies of state-sponsored homophobia.

References

- About-Chadi, T., Mitteregger, R., & Mudde, C. (2021). *Left Behind by the Working Class? Social Democracy's Electoral Crisis and the Rise of the Radical Right*. Friedrich-Ebert-Stiftung. Berlin. <https://www.fes.de/en/themenportal-demokratie-engagement-rechtsstaat-kommunalpolitik/article-page-flight-migration-integration/abandoned-by-the-working-class>
- About-Chadi, T., & Wagner, M. (2019). The electoral appeal of party strategies in postindustrial societies: When can the mainstream left succeed? *The Journal of Politics*, 81(4), 1405–1419. <https://doi.org/10.1086/704436>
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, 111(12), 4088–4118. <https://doi.org/10.1257/aer.20190159>
- Associated Press. (2019). Party leader calls LGBT rights an imported threat to Poland [newspaper]. *AP News*. <https://apnews.com/general-news-6cb144f6838e409c828cb32a19834a2a>
- Athey, S., Bayati, M., Doudchenko, N., Imbens, G. W., & Khosravi, K. (2021). Matrix Completion Methods for Causal Panel Data Models. *Journal of the American Statistical Association*, 116(536), 1716–1730. <https://doi.org/10.1080/01621459.2021.1891924>
- Ayoub, P. M. (2016). *When States Come Out: Europe's Sexual Minorities and the Politics of Visibility*. Cambridge University Press.
- Ayoub, P. M. (2020). Attacking LGBT life helped the right-wing polish president win reelection — barely [newspaper]. *Washington Post*. <https://www.washingtonpost.com/politics/2020/07/21/attacking-lgbt-life-helped-right-wing-polish-president-win-reelection-barely/>
- Ayoub, P. M., & Kollman, K. (2021). (Same)-sex in the city: Urbanisation and LGBTI rights expansion. *European Journal of Political Research*, 60(3), 603–624. <https://doi.org/10.1111/1475-6765.12418>
- Ayoub, P. M., & Stoeckl, K. (2024). The Global Resistance to LGBTIQ Rights. *Journal of Democracy*, 35(1), 59–73. <https://doi.org/10.1353/jod.2024.a915349>
- Bogatyrev, K., & Bogusz, H. (2025). On the verge of progress? LGBTQ+ politics in Poland after the 2023 elections. *European Journal of Politics and Gender*, 8(1), 242–248. <https://doi.org/10.1332/25151088Y2024D000000024>
- Bosia, M. J., & Weiss, M. L. (2013). Political Homophobia in Comparative Perspective. In M. L. Weiss & M. J. Bosia (Eds.), *Global Homophobia: States, Movements, and the Politics of Oppression*. University of Illinois Press. <https://doi.org/10.5406/illinois/9780252037726.003.0001>

- Bucholc, M. (2022). The anti-LGBTIQ campaign in Poland: The established, the outsiders, and the legal performance of exclusion. *Law & Policy*, 44(1), 4–22. <https://doi.org/10.1111/lapo.12183>
- Camp, B. J. (2008). Mobilizing the Base and Embarrassing the Opposition: Defense of Marriage Referenda and Cross-Cutting Electoral Cleavages. *Sociological Perspectives*, 51(4), 713–733. <https://doi.org/10.1525/sop.2008.51.4.713>
- Campbell, D. E., & Monson, J. Q. (2008). The Religion Card: Gay Marriage and the 2004 Presidential Election. *Public Opinion Quarterly*, 72(3), 399–419. <https://doi.org/10.1093/poq/nfn032>
- Charnysh, V. (2024). *Uprooted: How post-WWII Population Transfers Remade Europe*. Cambridge University Press. <https://doi.org/10.1017/9781009441995>
- Corrales, J. (2019). The Expansion of LGBT Rights in Latin America and the Backlash. In M. J. Bosia, S. M. McEvoy, & M. Rahman (Eds.), *The Oxford Handbook of Global LGBT and Sexual Diversity Politics* (pp. 185–200). Oxford University Press.
- Donovan, T., Tolbert, C. J., & Smith, D. A. (2008). Priming Presidential Votes by Direct Democracy. *The Journal of Politics*, 70(4), 1217–1231. <https://doi.org/10.1017/S0022381608081164>
- Düpont, N., & Rachuj, M. (2022). The ties that bind: Text similarities and conditional diffusion among parties. *British Journal of Political Science*, 52(2), 613–630. <https://doi.org/10.1017/S0007123420000617>
- European Parliament. (2020, September 17). *Sakharov Prize 2020: The nominees*. <https://www.europarl.europa.eu/topics/en/article/20200917STO87301/sakharov-prize-2020-the-nominees>
- Flores, A. R., & Barclay, S. (2016). Backlash, Consensus, Legitimacy, or Polarization: The Effect of Same-Sex Marriage Policy on Mass Attitudes. *Political Research Quarterly*, 69(1), 43–56. <https://doi.org/10.1177/1065912915621175>
- Garretson, J. J. (2014). Changing with the Times: The Spillover Effects of Same-Sex Marriage Ballot Measures on Presidential Elections. *Political Research Quarterly*, 67(2), 280–292. <https://doi.org/10.1177/1065912914521897>
- Gibson, E. L. (2005). Boundary Control: Subnational Authoritarianism in Democratic Countries. *World Politics*, 58(1), 101–132. <https://doi.org/10.1353/wp.2006.0018>
- Graff, A., & Korolczuk, E. (2022). “Worse Than Communism and Nazism Put Together”: Poland’s Anti-Gender Campaigns in a Comparative Perspective. In *Anti-Gender Politics in the Populist Moment* (pp. 67–91). Routledge.
- Grosfeld, I., & Zhuravskaya, E. (2015). Cultural vs. economic legacies of empires: Evidence from the partition of Poland. *Journal of Comparative Economics*, 43(1), 55–75. <https://doi.org/10.1016/j.jce.2014.11.004>
- Grumbach, J. M. (2023). Laboratories of Democratic Backsliding. *American Political Science Review*, 117(3), 967–984. <https://doi.org/10.1017/S0003055422000934>
- Gwiazda, A. (2023). Gender Ideologies and Polish Political Parties. *Government and Opposition*, 58(4), 641–660. <https://doi.org/10.1017/gov.2021.57>

- Hennig, A. (2023). The Diverse Catholic Right in Poland. In G. Lo Mascolo (Ed.), *The Christian Right in Europe* (1st ed., pp. 75–104). transcript Verlag. <https://doi.org/10.14361/9783839460382-005>
- Hillygus, D. S., & Shields, T. G. (2005). Moral Issues and Voter Decision Making in the 2004 Presidential Election. *PS: Political Science and Politics*, 38(2), 201–209. <https://doi.org/10.1017/S1049096505056301>
- ILGA-Europe. (2019). *Annual Review of the Human Rights Situation of Lesbian, Gay, Bisexual, Trans and Intersex People*. <https://new.ilga-europe.org/annualreview/2019>
- Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M., & Vachudova, M. A. (2022). Chapel Hill Expert Survey trend file, 1999–2019. *Electoral Studies*, 75, 102420. <https://doi.org/10.1016/j.electstud.2021.102420>
- Lewis, G. B. (2005). Same-Sex Marriage and the 2004 Presidential Election. *PS: Political Science & Politics*, 38(2), 195–199. <https://doi.org/10.1017/S1049096505056295>
- Magni, G., & Reynolds, A. (2023). Candidate Identity and Campaign Priming: Analyzing Voter Support for Pete Buttigieg’s Presidential Run as an Openly Gay Man. *Political Research Quarterly*, 10659129231194325. <https://doi.org/10.1177/10659129231194325>
- Markowski, R., Czesnik, M., Stanley, B., Wenzel, M., Zerkowska-Balas, M., & Zagórski, P. (2023, March 30). *Polish national election study 2019 (PNES 2019)*. Harvard Dataverse. <https://doi.org/10.7910/DVN/94B2AU>
- Meyerhoefer, C. D., Xue, B., & Poznańska, A. (2025). Implications of the Decline in LGBT Rights for Population Mental Health: Evidence from Polish “LGBT-free zones”. *Journal of Health Economics*, 102973. <https://doi.org/10.1016/j.jhealeco.2025.102973>
- O’Dwyer, C. (2012). Does the EU Help or Hinder Gay-Rights Movements in Post-Communist Europe? The Case of Poland. *East European Politics*, 28(4), 332–352. <https://doi.org/10.1080/21599165.2012.721094>
- O’Dwyer, C., & Stenberg, M. (2022). Local-Level Democratic Backsliding? The Consolidation of Aspiring Dominant-Party Regimes in Hungary and Poland. *Government and Opposition*, 57(3), 508–531. <https://doi.org/10.1017/gov.2021.12>
- Page, D., Ayoub, P., Arranz, C., Montes, M., & Paulin, T. (2022). Reassessing the relationship between homophobia and political participation. *European Journal of Political Research*, 61(4), 1165–1178. <https://doi.org/10.1111/1475-6765.12519>
- Pająk, P., & Gawron, J. (2020). *Atlas of hate*. <https://atlasnienawisci.pl/>
- Parliament, E. (2019, December 18). *Resolution on Public Discrimination and Hate Speech Against LGBTI People, Including LGBTI Free Zones* (2019/2933(RSP)). https://www.europarl.europa.eu/doceo/document/TA-9-2019-0101_EN.html
- Payne, L. A., & Santos, A. A. d. S. (2020). The Right-Wing Backlash in Brazil and Beyond. *Politics & Gender*, 16(1), 32–38. <https://doi.org/10.1017/S1743923X20000057>

- Ptak, A. (2022, December 22). *Activists mapping Poland's anti-LGBT zones win in court against county that sued over "Atlas of Hate"*. Notes from Poland. <https://notesfrompoland.com/2022/12/22/activists-mapping-polands-anti-lgbt-zones-win-in-court-against-county-that-sued-over-atlas-of-hate/>
- Rafałowski, W. (2023). Proposing Anti-LGBTQIAP Resolutions at the Municipal Level in Poland: Meeting the Social Demand or Making Use of the Available Resources? *East European Politics and Societies*, 37(2), 627–655. <https://doi.org/10.1177/08883254221082133>
- Reuters. (2023). Erdogan calls opposition 'pro-LGBT' at election rally [newspaper]. *Reuters: Middle East*. <https://www.reuters.com/world/middle-east/erdogan-calls-opposition-pro-lgbt-election-rally-2023-05-07/>
- Smith, A. E., & Boas, T. C. (2024). Religion, Sexuality Politics, and the Transformation of Latin American Electorates. *British Journal of Political Science*, 54(3), 816–835. <https://doi.org/10.1017/S0007123423000613>
- Smith, D. A., DeSantis, M., & Kassel, J. (2006). Same-Sex Marriage Ballot Measures and the 2004 Presidential Election. *State & Local Government Review*, 38(2), 78–91. <https://doi.org/10.1177/0160323X0603800202>
- Stenberg, M., & O'Dwyer, C. (2023). *When do local governments discriminate? Lessons from the spread of "LGBT free zones" in Poland*. <https://doi.org/10.31235/osf.io/8e7wh>
- Świętokrzyskie Voivodeship Assembly. (2019, August 26). *Resolution XII/168/19*. <https://bip.sejmik.kielce.pl/55-uchwaly-sejmiku.html>
- Turnbull-Dugarte, S. J., & Ortega, A. L. (2024). Instrumentally Inclusive: The Political Psychology of Homonationalism. *American Political Science Review*, 118(3), 1360–1378. <https://doi.org/10.1017/S0003055423000849>
- Wanat, Z. (2020). Polish towns pay a steep price for anti-LGBTQ views [magazine]. *Politico*. <https://www.politico.eu/article/poland-lgbtq-steep-price/>

2.A Appendix

2.A.1 Parliamentary Election 2019: Anti-LGBTQ Campaign



Figure 2.8. Example for anti-LGBTQ campaign by the then-president of the Polish Senate, Karczewski, Stanisław (@StKarczewski) in the run-up to the 2019 parliamentary election. English translation: “Chairman PiS J. Kaczyński in Rzeszów: We say No! to the attack on children. We will not be intimidated. We will defend the Polish family” Twitter/X, March 10, 2019, 6:53am.

2.A.2 Party Positions on LGBTQ Issues in 2019

Party / Alliance	Ideology	CHES 2019 Social Lifestyle (0-10)	V-Party 2019 LGBT Equality (0-4)
<i>Government Party</i>			
Law and Justice (PiS)	radical right	8.7 (strongly opposed)	0.4 (strongly opposed)
<i>Opposition Parties</i>			
Civic Coalition (KO)	centrist-liberal	4.1 (moderate)	2.5 (moderate)
Polish Coalition (KP)	centrist- conservative	6.9 (opposed)	1.4 (opposed)
The Left (Lewica)	left	1.4 (strongly in favor)	3.0 (in favor)
Confederation	extreme right	9.7 (strongly opposed)	0.4 (strongly opposed)

Note: The third column represents party scores for item 25 “Position on social lifestyle policy (e.g., rights for homosexuals, gender equality)” from the Chapel Hill Expert Survey 2019 (Jolly et al., 2022). The fourth column represents party scores on the original scale for item 3.2.8 “What is this party’s position toward social equality for the lesbian, gay, bisexual, and transgender (LGBT) community?” for 2019 from the Varieties of Party Identity and Organization (V-Party v2) dataset (Düpont & Rachuj, 2022). Both datasets are expert-coded. For party alliances, scores were obtained by weighting party positions present in the data by the party’s share in the alliance.

Table 2.1. Party Positions in the 2019 Polish Parliamentary Election

2.A.3 Synthetic Control Weights

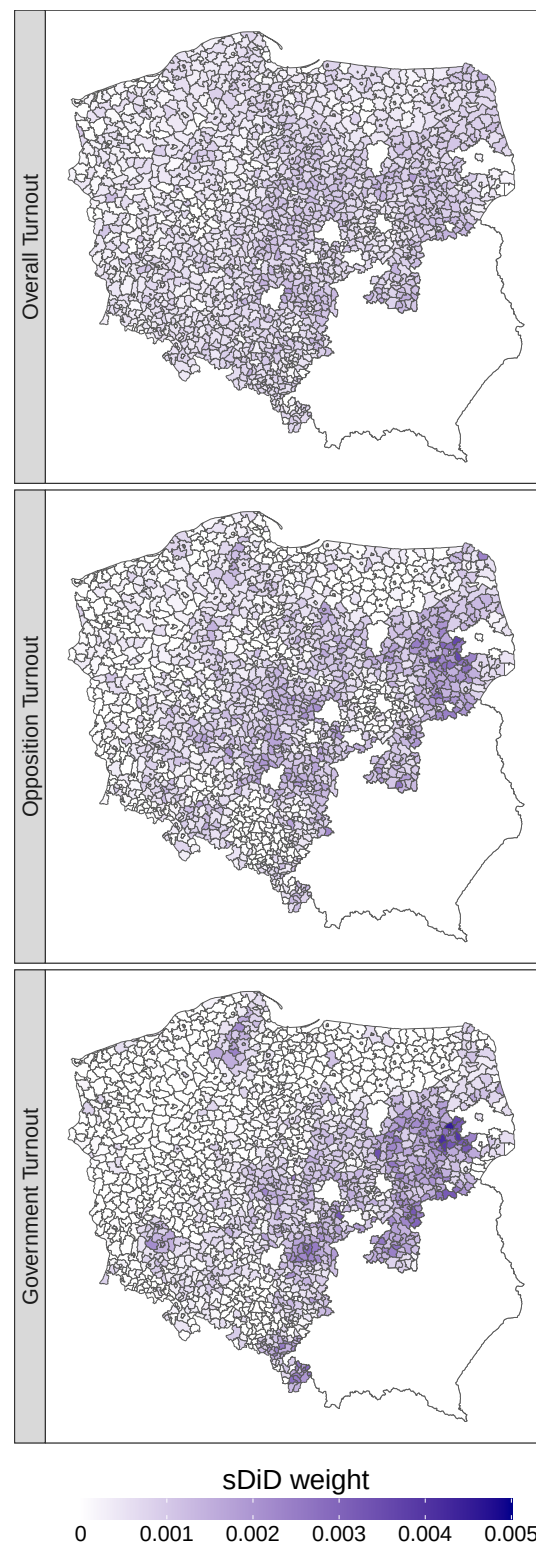


Figure 2.9. Geographic distribution of synthetic control weights for overall, opposition, and government turnout.

2.A.4 Restricted Sample

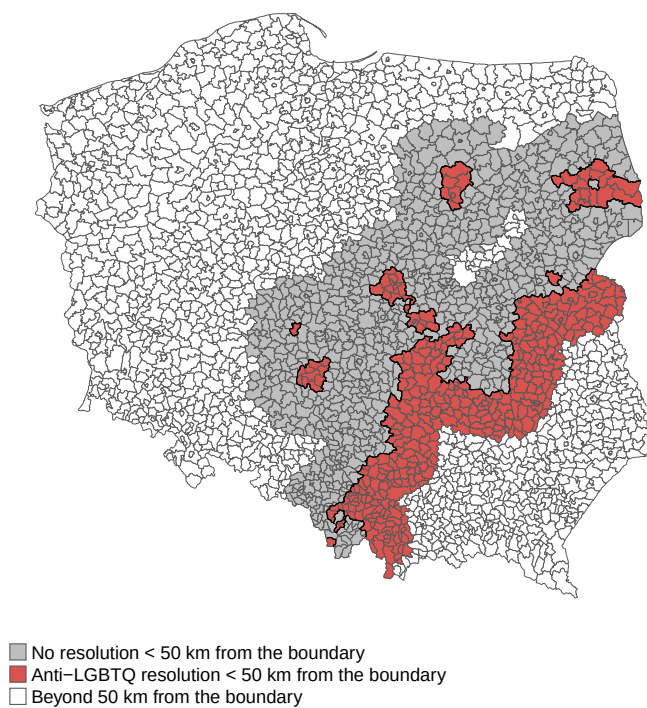


Figure 2.10. Geographic distribution of municipalities within 50 km distance from treatment-control boundaries.

2.A.5 Full sDiD Output

Outcome	Sample	Estimate (SE)	N
Overall Turnout	Full Sample	-1.69 (0.11)	14856
	< 50km Sample	-0.73 (0.15)	6594
Opposition Turnout	Full Sample	-1.28 (0.14)	14856
	< 50km Sample	-0.89 (0.19)	6594
Government Turnout	Full Sample	0.41 (0.18)	14856
	< 50km Sample	0.67 (0.24)	6594

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.2. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout by sample.

2.A.6 Effects on Vote Shares and Turnout by Party

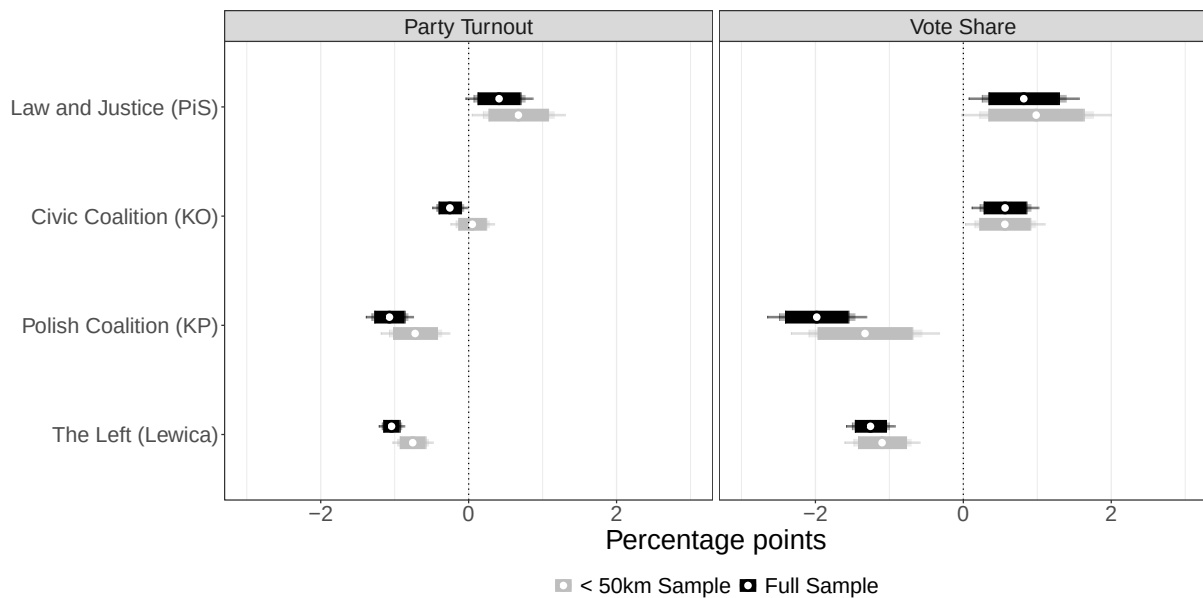


Figure 2.11. Effect of anti-LGBTQ resolutions on party-specific turnout and vote shares in the 2019 parliamentary election by sample. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table 2.3.

Party	Sample	Estimate (SE)	N
<i>Party Turnout</i>			
Law and Justice (PiS)	Full Sample	0.41 (0.18)	14856
	< 50km Sample	0.67 (0.24)	6594
Civic Coalition (KO)	Full Sample	-0.25 (0.09)	14856
	< 50km Sample	0.05 (0.11)	6594
Polish Coalition (KP)	Full Sample	-1.07 (0.12)	14856
	< 50km Sample	-0.72 (0.18)	6594
The Left (Lewica)	Full Sample	-1.04 (0.07)	14856
	< 50km Sample	-0.76 (0.11)	6594
<i>Vote Share</i>			
Law and Justice (PiS)	Full Sample	0.82 (0.29)	14856
	< 50km Sample	0.98 (0.39)	6594
Civic Coalition (KO)	Full Sample	0.56 (0.18)	14856
	< 50km Sample	0.56 (0.21)	6594
Polish Coalition (KP)	Full Sample	-1.98 (0.26)	14856
	< 50km Sample	-1.33 (0.39)	6594
The Left (Lewica)	Full Sample	-1.26 (0.13)	14856
	< 50km Sample	-1.10 (0.20)	6594

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.3. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on party-specific turnout and vote shares by sample.

2.A.7 Varying Treatment Intensity

2.A.7.1 ATT by Sum of Government Levels

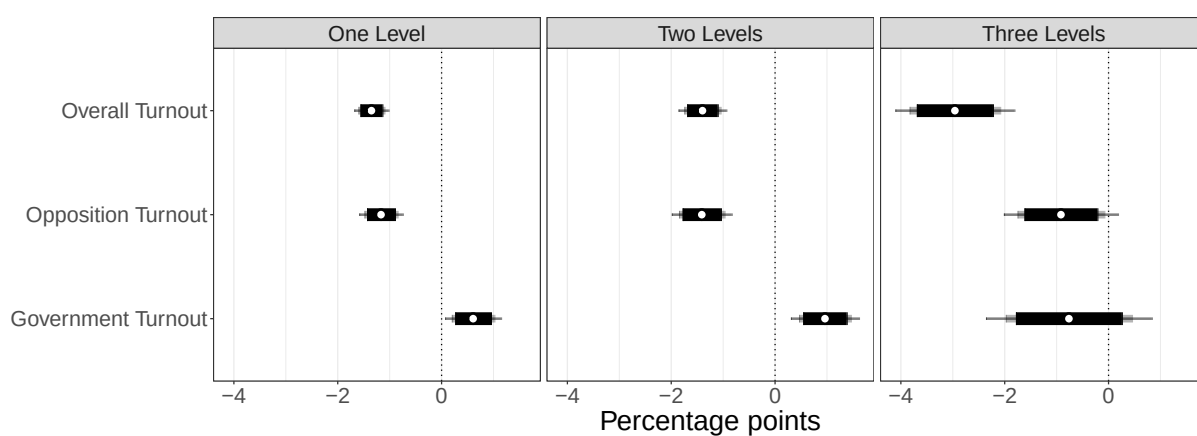


Figure 2.12. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout in the 2019 parliamentary election by sum of treatment levels. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table 2.4.

Outcome	Sum Level	Estimate (SE)	N
Overall Turnout	One Level	-1.35	12114
		(0.13)	
	Two Levels	-1.40	11034
		(0.18)	
	Three Levels	-2.96	9576
		(0.45)	
Opposition Turnout	One Level	-1.17	12114
		(0.16)	
	Two Levels	-1.41	11034
		(0.22)	
	Three Levels	-0.92	9576
		(0.43)	
Government Turnout	One Level	0.61	12114
		(0.21)	
	Two Levels	0.96	11034
		(0.25)	
	Three Levels	-0.76	9576
		(0.62)	

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.4. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout by sum of treatment levels.

2.A.7.2 ATT by Sum of Resolutions

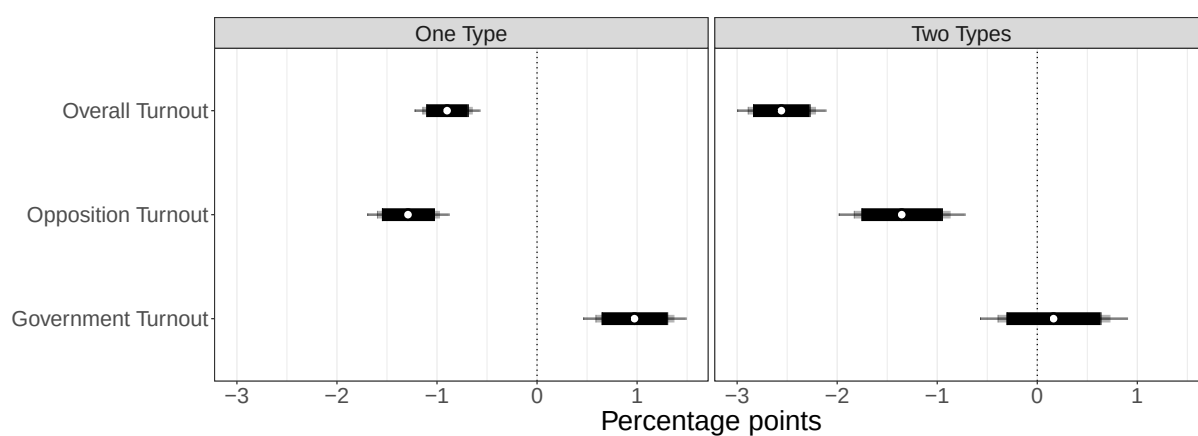


Figure 2.13. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout in the 2019 parliamentary election by sum of treatment resolutions. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table 2.5.

Outcome	Sum Type	Estimate (SE)	N
Overall Turnout	One Type	-0.90 (0.13)	12402
	Two Types	-2.56 (0.17)	10752
Opposition Turnout	One Type	-1.29 (0.16)	12402
	Two Types	-1.35 (0.24)	10752
Government Turnout	One Type	0.97 (0.20)	12402
	Two Types	0.16 (0.29)	10752

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., [2021](#)).

Table 2.5. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout by sum of treatment resolutions.

2.A.8 Alternative Treatment Definitions

2.A.8.1 Geographic Distribution of Treatment by Government-Level

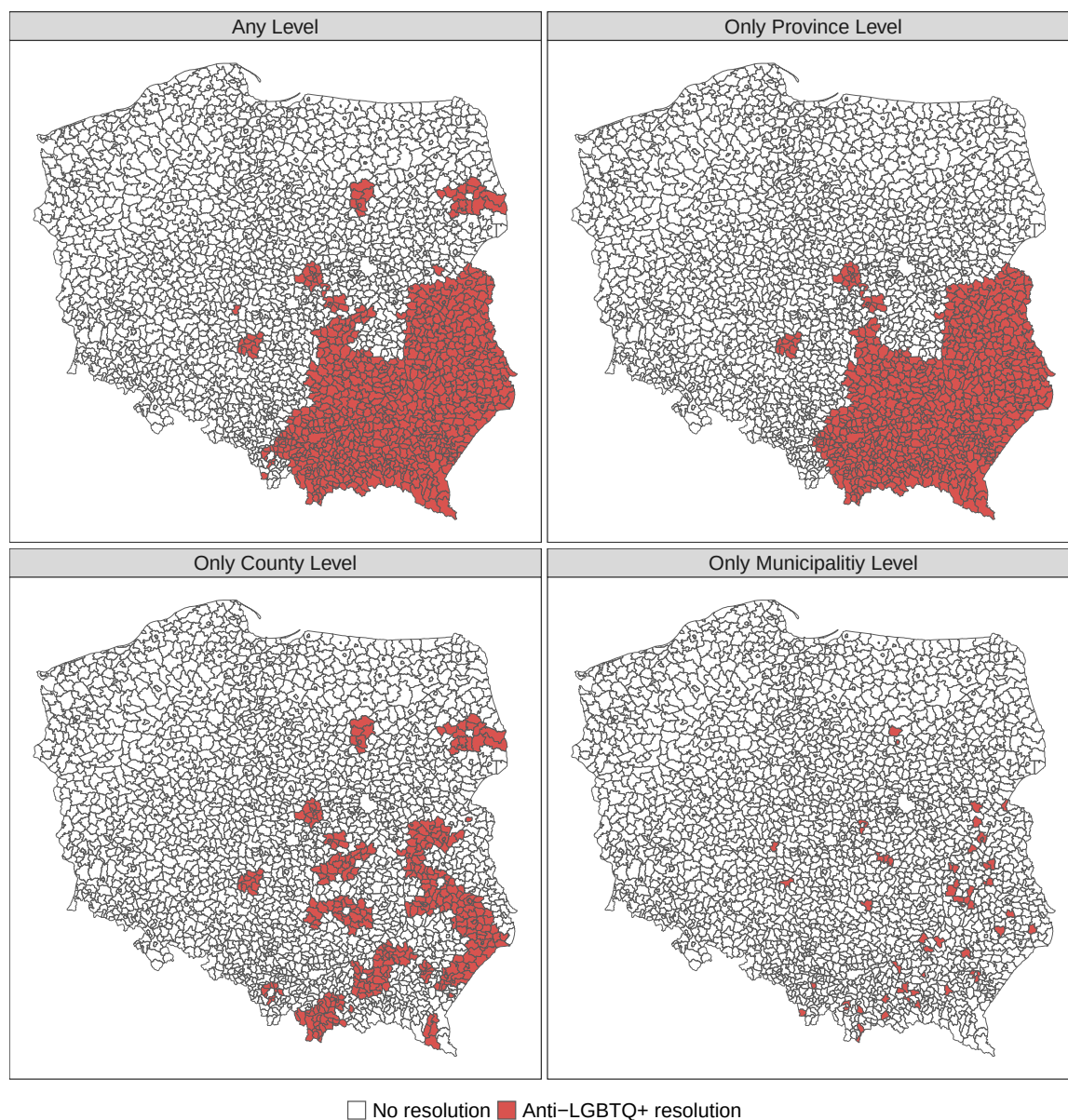


Figure 2.14. Geographic distribution of municipalities with anti-LGBTQ resolutions in effect prior to the 2019 parliamentary election by government level passing the resolution. A) any government-level, B) only province-level, C) only county-level, D) only municipality-level.

2.A.8.2 ATT by Government-Level

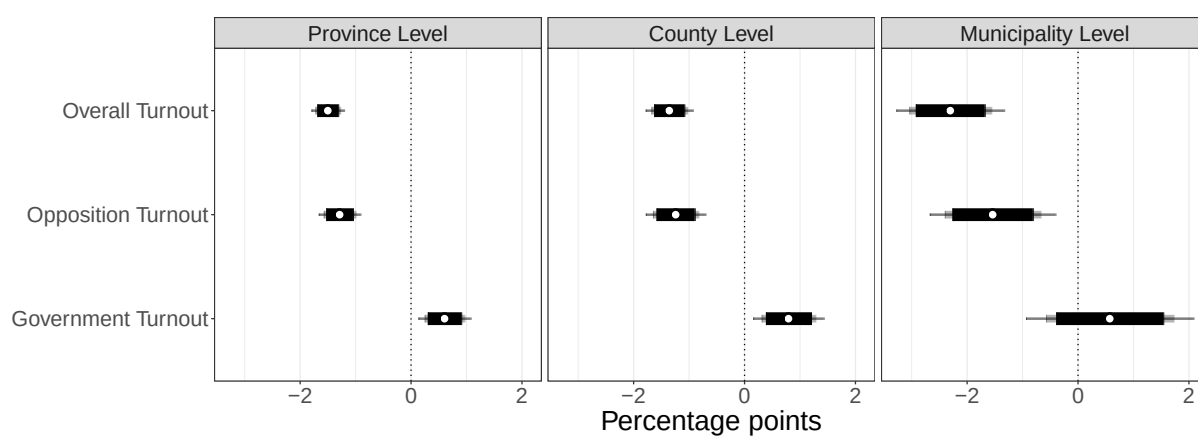


Figure 2.15. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout in the 2019 parliamentary election, separated by government unit passing the resolution. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table 2.6.

Outcome	Level	Estimate (SE)	N
Overall Turnout	Province Level	-1.50 (0.11)	13578
	County Level	-1.36 (0.17)	11316
	Municipality Level	-2.30 (0.38)	9738
Opposition Turnout	Province Level	-1.29 (0.15)	13578
	County Level	-1.24 (0.21)	11316
	Municipality Level	-1.54 (0.44)	9738
Government Turnout	Province Level	0.60 (0.18)	13578
	County Level	0.79 (0.25)	11316
	Municipality Level	0.57 (0.59)	9738

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.6. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, separated by government unit passing the resolution.

2.A.8.3 ATT by Resolution Type

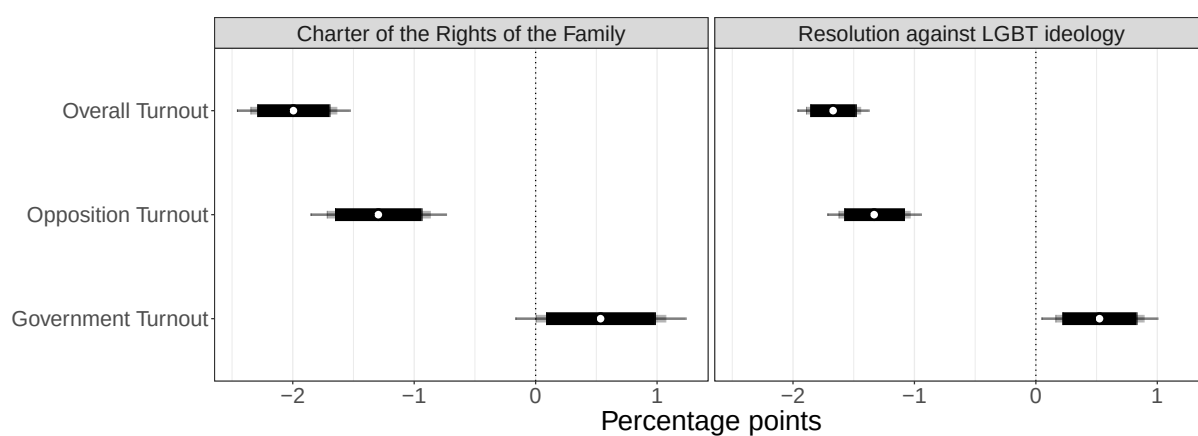


Figure 2.16. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout in the 2019 parliamentary election, separated by type of resolution. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table 2.7.

Outcome	Resolution Type	Estimate (SE)	N
Overall Turnout	Charter of the Rights of the Family	-1.99 (0.18)	12066
	Resolution against LGBT ideology	-1.67 (0.11)	13398
Opposition Turnout	Charter of the Rights of the Family	-1.30 (0.22)	12066
	Resolution against LGBT ideology	-1.33 (0.15)	13398
Government Turnout	Charter of the Rights of the Family	0.53 (0.27)	12066
	Resolution against LGBT ideology	0.52 (0.19)	13398

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.7. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, separated by type of resolution.

2.A.9 Alternative Control Sample

2.A.9.1 Never-Treated Synthetic Control Group

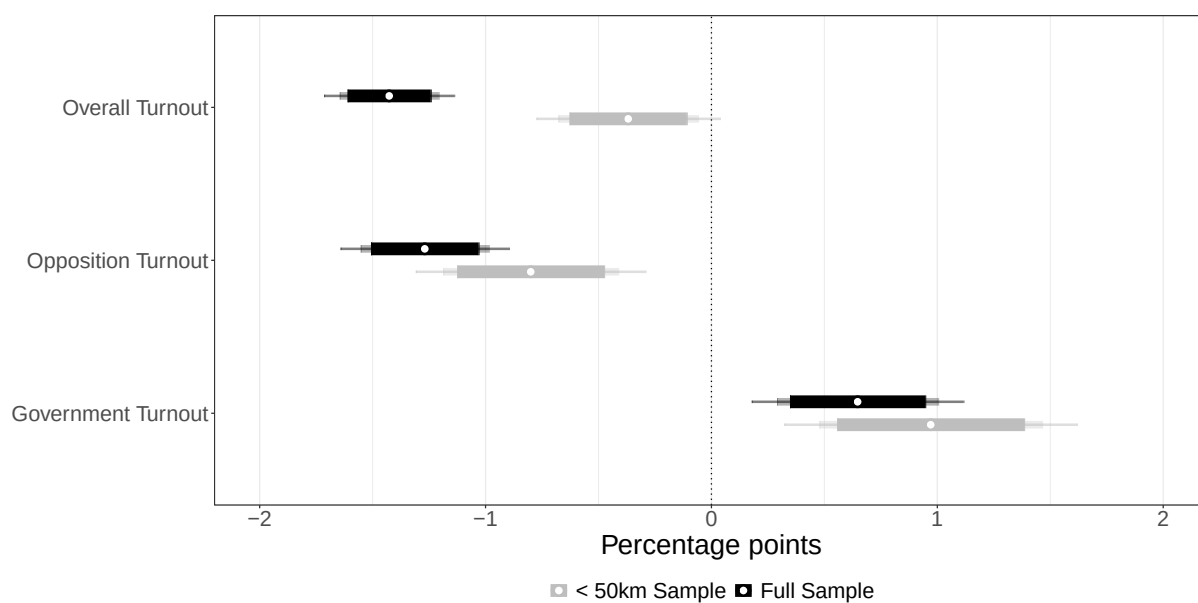


Figure 2.17. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout. We report point estimates from synthetic difference-in-differences estimates along 95% confidence intervals. The synthetic control-generating process draws on never-treated units. Full output in Appendix Table [2.8](#).

Outcome	Sample	Estimate (SE)	N
Overall Turnout	Full Sample	-1.43 (0.11)	13872
	< 50km Sample	-0.37 (0.16)	5658
Opposition Turnout	Full Sample	-1.27 (0.14)	13872
	< 50km Sample	-0.80 (0.20)	5658
Government Turnout	Full Sample	0.65 (0.18)	13872
	< 50km Sample	0.97 (0.25)	5658

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.8. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout using the never-treated control unit pool.

2.A.10 Additional Robustness Checks

2.A.10.1 Alternative Estimators

Outcome	Sample	DiD	MC	sDiD
Overall Turnout	Full Sample	-1.09	-1.29	-1.69
		(0.15)	(0.11)	(0.1)
	< 50km Sample	-0.56	-0.59	-0.73
		(0.24)	(0.16)	(0.15)
Opposition Turnout	Full Sample	-5.68	-2.24	-1.28
		(0.24)	(0.13)	(0.1)
	< 50km Sample	-2.26	-1.12	-0.89
		(0.41)	(0.19)	(0.16)
Government Turnout	Full Sample	4.67	2.19	0.41
		(0.24)	(0.18)	(0.13)
	< 50km Sample	1.77	1.14	0.67
		(0.38)	(0.28)	(0.22)

Note: Placebo standard errors (Arkhangelsky et al., 2021) in parentheses.

N = 14,856 for all models. DiD stands for “difference-in-differences”, MC stands for “matrix completion”, sDiD stands for “synthetic difference-in-differences”.

Table 2.9. Estimated effects of anti-LGBTQ resolutions on overall, opposition, and government turnout in treated municipalities in the 2019 parliamentary election by estimator.

2.A.10.2 Excluding Municipalities with over 300,000 Inhabitants

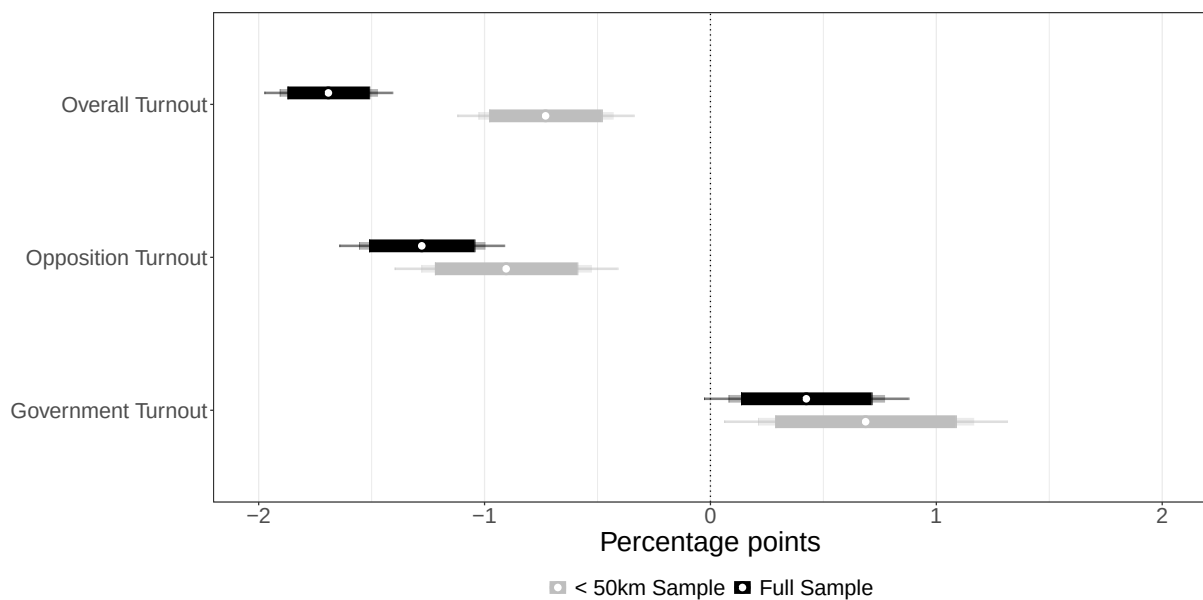


Figure 2.18. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, excluding the ten largest municipalities with over 300,000 inhabitants. We report point estimates from synthetic difference-in-differences estimates along 90%, 95%, and 99% confidence intervals. Full output in Appendix Table [2.10](#).

Outcome	Sample	Estimate (SE)	N
Overall Turnout	Full Sample	-1.69 (0.11)	14820
	< 50km Sample	-0.73 (0.15)	6576
Opposition Turnout	Full Sample	-1.28 (0.14)	14820
	< 50km Sample	-0.90 (0.19)	6576
Government Turnout	Full Sample	0.42 (0.18)	14820
	< 50km Sample	0.69 (0.24)	6576

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.10. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, excluding the ten largest municipalities with over 300,000 inhabitants.

2.A.10.3 Excluding Municipalities Proposing but not Adopting Resolutions

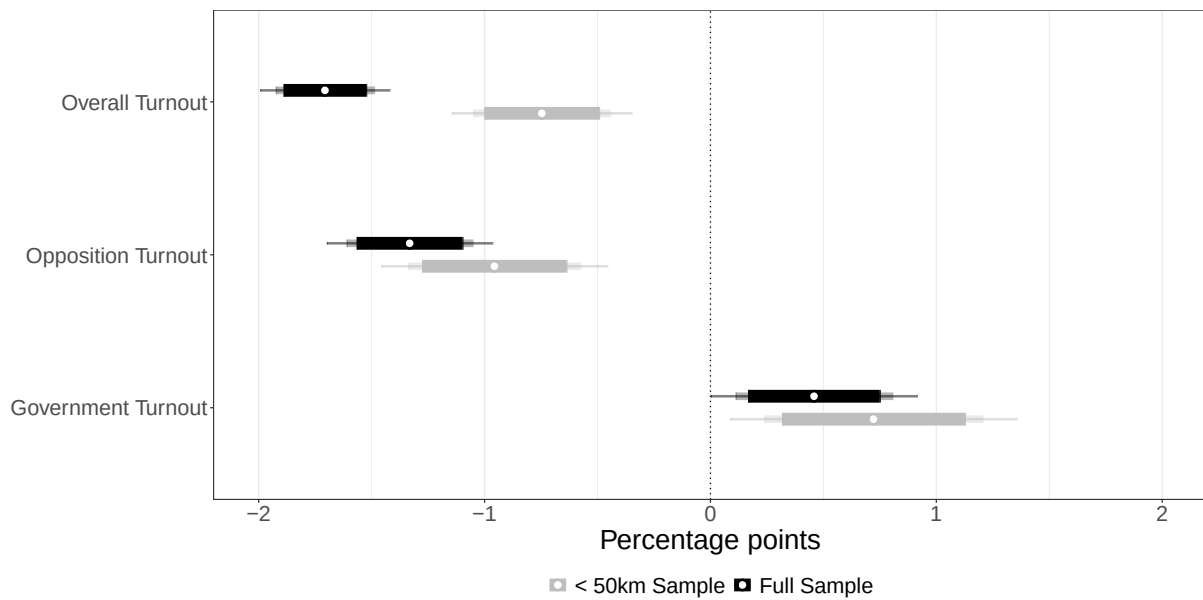


Figure 2.19. Effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, excluding municipalities that proposed but did not adopt resolutions (i.e., discussed, rejected or dismissed proposals). We report point estimates from synthetic difference-in-differences estimates along 95% confidence intervals. Full output in Appendix Table 2.11.

Outcome	Sample	Estimate (SE)	N
Overall Turnout	Full Sample	-1.71 (0.11)	14466
	< 50km Sample	-0.75 (0.15)	6396
Opposition Turnout	Full Sample	-1.33 (0.14)	14466
	< 50km Sample	-0.96 (0.19)	6396
Government Turnout	Full Sample	0.46 (0.18)	14466
	< 50km Sample	0.72 (0.25)	6396

Note: Standard errors (in parentheses) were estimated using jackknife variance estimation (Arkhangelsky et al., 2021).

Table 2.11. Synthetic difference-in-differences estimates for the effect of anti-LGBTQ resolutions on overall, opposition, and government turnout, excluding municipalities that proposed but did not adopt resolutions (i.e., discussed, rejected or dismissed proposals).

2.A.11 Municipality Count by Treatment Status & Sample

Sample	Never Treated	Treated before 2019 election	Treated after 2019 election	Total
Full Sample	1572	741	164	2477
One Level	”	448	152	600
Two Levels	”	268	11	279
Three Levels	”	25	1	26
One Type	”	448	152	600
Two Types	”	268	11	279
Province Level	”	692	142	834
County Level	”	315	29	344
Municipality Level	”	52	6	58
Resolution against LGBT ideology	”	662	0	662
Charter of the Rights of the Family	”	276	164	440
No municipalities > 300,000	1567	740	163	2470
No proposing municipalities	1521	728	163	2412
< 50km Sample	586	357	156	1099
One Level	”	218	144	362
Two Levels	”	126	11	137
Three Levels	”	13	1	14
One Type	”	218	144	362
Two Types	”	126	11	137
Province Level	”	308	142	450
County Level	”	173	24	197
Municipality Level	”	28	3	31
Resolution against LGBT ideology	”	278	0	278
Charter of the Rights of the Family	”	85	156	241
No municipalities > 300,000	585	356	155	1096
No proposing municipalities	564	347	155	1066

Note: The column “Treated after the 2019 election” only counts the newly treated municipalities, i.e. those not treated before the election by any resolution type adopted at any administrative level.

Chapter 3

Reciprocal Shifts: Synthetic Control Methods for Proportions¹

Synthetic control methods are widely used for causal inference in case studies and panel data settings, often applied to model counterfactuals for proportional outcomes. However, conventional synthetic control methods are designed for univariate outcomes, leading researchers to model counterfactuals for each proportion separately. This study introduces an extension to synthetic control methods to simultaneously handle multivariate outcomes in compositional data. The approach establishes constant control comparisons by using the same weights for each proportion, improving comparability while adhering to treatment constraints. Results from a simulation study and an empirical application of the method to the study from chapter 2 and another study on climate policy effects underscore the benefits of accounting for the interplay of proportional outcomes. This advancement extends the validity and reliability of synthetic control estimates to common outcomes in political science.

¹This chapter is based on a working paper co-authored with Lukas Stoetzer, with [an empirical application to Poland's case](#) added by me for the dissertation.. I am grateful to Lukas, as well as to Tom Pepinsky, Paul Kellstedt and the participants of the PhD seminar at Bocconi University for their helpful comments, and my advisor Catherine E. De Vries for encouraging my effort to work on political methodology and supporting it with funding from the European Research Council (ERC Grant Agreement 864687).

3.1 Introduction

Political and other social scientists are often interested in studying the effects of policies, events and interventions. For this purpose, causal inference methods such as difference-in-differences (DID) have proven to be incredibly helpful and have gained widespread use in the analysis of time-series cross-sectional data, or long panel data (Arkhangelsky & Imbens, 2024; Liu et al., 2024; Xu, 2024). However, DID-type estimates commonly assume parallel trends between treatment and control units in the absence of treatment, which proves difficult to evaluate and is often implausible (see e.g. Chiu et al., 2024; Hassell & Holbein, 2024; Roth, 2022). Following the premise that a convex weighted combination of control units can better replicate the counterfactual trajectory for the treated (Abadie et al., 2015), synthetic controls have gained recognition as a path-breaking innovation for causal inference with time-series cross-sectional data (Xu, 2024). Synthetic control methods (SCMs)² found a variety of applications to modeling univariate outcomes in political science, economics, and other social sciences.³

In many applications, however, political scientists deal with multivariate, compositional data in the form of proportions. As Katz and King noted already 25 years ago: “Compositional data are also common in political science and economics (as in multiparty voting data, the allocation of ministerial portfolios among political parties, trade flows or international conflict directed from each nation to several others, or proportions of budget expenditures in each of several categories)” (Katz & King, 1999). In compositional data, each proportion is bound between 0 and 1, and all proportions sum up to 1. Due to these simple yet important properties, compositional data often require separate methods (Jackson, 2002; Lipsmeyer et al., 2019; Philips et al., 2016; Tomz et al., 2002). As we

²Hereinafter, we use the term “synthetic control methods” or simply “synthetic controls” to denote the entire family of methods, thus not limiting our discussion to the original method but also including recent advances such as generalized (Xu, 2017), penalized (Abadie & L’Hour, 2021) and augmented (Ben-Michael et al., 2021) synthetic controls, as well as synthetic DID (Arkhangelsky et al., 2021).

³This includes such outcomes as single-party vote shares and turnout (Becher et al., 2023; de Kadt et al., 2024; Hilbig & Riaz, 2024; Lal et al., 2024; Thompson, 2023), policy adoption (Gilens et al., 2021; Meserve & Pemstein, 2018), welfare outcomes (Eibl & Hertog, 2024), crime (Alcocer, 2024; Alrababa’h et al., 2021; Pinotti, 2015; Trejo & Nieto-Matiz, 2023), GDP (de Kadt & Wittels, 2019; Funke et al., 2023; Pinotti, 2015) and power personalization (Timoneda et al., 2023).

show further on, SCMs are no exception.

When the primary outcomes of interest are proportions of the same whole, the separate application of SCMs to the different proportions violates – what we will call – the constant control comparison for compositional outcomes. Each synthetic control application generates its own set of weights, which results in different control groups for each proportion. This can result in treatment effect estimates that violate an essential sum constraint. In compositional data, the sum of changes should generally equal zero because an increase in one proportion necessarily implies a decrease in another.

We expand the synthetic control methods framework to align with the constant control comparison for compositional outcomes. This extension introduces a method for estimating models that utilize the same control weights across different proportions. We demonstrate that this model adheres to the sum constraint on treatment effects, thereby offering a more accurate interpretation of treatment effects for proportional outcomes. A freely available R package, `propsdid`, implements the estimation.

A simulation study reveals that the synthetic control methods for proportions reduce bias compared to DID methods (especially when the parallel trends assumption fails) and are more reliable, compared to the separate application of SCMs that fail to meet the sum constraint of average treatment effects. The application of our method to a recent study in the American Political Science Review (Bolet et al., 2024) further demonstrates the benefits of analyzing voting data as multivariate proportional outcomes. The authors argue that in the 2019 Spanish general election, the PSOE government gained votes in former coal areas by supporting local communities in their transition away from carbon-intensive industries. Applying the synthetic control methods for proportions shows that not only PSOE but also their main conservative competitors gained in these areas, which challenges the main findings of the replicated study.

Our paper makes a set of methodological contributions. We bridge the older and more established literature on compositional data (Aitchison, 1982; Jackson, 2002; Katz & King, 1999; Lipsmeyer et al., 2019; Philips et al., 2016; Tomz et al., 2002) with the

younger and burgeoning field of causal inference in panel data settings (recently reviewed by Arkhangelsky & Imbens, 2024; Roth et al., 2023; Xu, 2024). We highlight the relevance of constant control comparison for modeling compositional outcomes in any time-series cross-sectional setup, regardless of whether SCMs are used. We address the shortcomings of existing synthetic control methods in handling compositional data and develop a novel approach to synthetic controls for proportions. Importantly, our approach can serve as a building block for both the original SC method (Abadie et al., 2010, 2015) and its advanced successors (Arkhangelsky et al., 2021; Ben-Michael et al., 2021; Xu, 2017).

The paper is organized as follows. In Section 3.2, we review the theoretical foundations and recent advances in synthetic control methods. Section 3.3 demonstrates the challenges of applying SCMs to compositional data in political and social science. We formally introduce our method extension for proportions in Section 3.4. Then, Section 3.5 presents a simulation study to evaluate the performance of our method compared to traditional approaches. In Section 3.6.1, we apply our method to the case study of the 2019 Spanish general election. We conclude with a discussion of the implications that our results provide for research on causal inference with compositional panel data.

3.2 Synthetic Control Methods: Current Approaches

Researchers often seek to evaluate the effects of policies, events, and interventions using panel data, i.e. repeated observations of the same units over time, when some units are exposed to these treatments during certain periods and not in others. The main difficulty is that these treatment changes are typically not random across units or time, which challenges the assumption of unconfoundedness, even when accounting for observed covariates. To address this challenge, difference-in-differences (DID) methods and two-way fixed effects regressions have become the dominant approaches to causal inference with panel data in political science and economics (Currie et al., 2020; Xu, 2024). Yet, synthetic control methods (SCMs) have emerged as a promising alternative, particularly for applications where the assumptions of DID are harder to justify.

These methods create a synthetic “untreated” version of the treated group by reweighting control units to match the pre-exposure trends of the treated. Thus, SCMs address the situations where the treatment and control trends are not parallel before exposure, indicating a likely violation of the (post-exposure) parallel trends assumption required by the DID methods (Xu, 2017). While the original synthetic control (SC) method (Abadie et al., 2010, 2015) was designed for small-sample single-unit case studies and did not allow for intercept shifts between the treated and control trajectories, recent methodological advancements (e.g., Arkhangelsky et al., 2021; Ben-Michael et al., 2021; Xu, 2017) have addressed those limitations and shown improved performance, both in small and large panel data settings.

For further discussion of SCMs in general, the synthetic difference in differences (SDID) approach by Arkhangelsky et al. (2021) provides a useful methodological framework, offering a generalization of both DID and SC and combining their strengths. The SDID method matches and reweights pre-exposure trends, reducing reliance on the parallel trends assumption, akin to SC. It also maintains the invariant properties of DID related to the additive unit-level shifts, allowing for valid inference across large panels. The reweighted post-treatment control outcomes in both methods then serve as the comparison for the estimation of average treatment effects (on the treated). For a formal description of SCMs following Arkhangelsky et al. (2021), please refer to the Appendix 3.A.1. Arkhangelsky et al. (2021) demonstrate both the theoretical robustness of their estimator and its empirical competitiveness with, or superiority to, traditional DID and SC methods in their respective typical applications.

Despite the recent methodological generalizations of SCMs, the outcomes modeled by these methods are univariate. Although current approaches allow for studying the effects on a particular outcome, often the outcomes of interest are multivariate and inextricably linked to one another. This is particularly true for compositional data, where proportions of the same whole are bound by a linear relationship. Thus, compositional outcome variables present distinct challenges for causal inference (Arnold et al., 2020).

However, synthetic control methods were not originally designed for multivariate outcomes, leaving their suitability for compositional data an open question. Researchers have, however, addressed the issue of multivariate outcomes in panel data more generally. A common strategy is to model outcomes jointly, relying on the joint distribution rather than treating each outcome independently (Mullahy, 2018; Negi & Wooldridge, 2024; Samartsidis et al., 2021). For analyzing the treatment effects on each outcome, scholars proposed to account for multiple testing or to jointly test for effects on any outcomes (Athey & Imbens, 2017; Robbins et al., 2017). For SCMs in particular, Sun et al. (2023) suggest optimizing the pre-treatment fit for concatenated standardized outcomes or the average of standardized outcomes.

Building upon this literature, our paper provides a statistical foundation for using SCMs with compositional data. In the following section, we discuss the issues that arise when researchers apply SCMs to proportions.

3.3 Compositional Effect Issues with Synthetic Control Estimates

3.3.1 A Motivating Example

Before discussing the framework in abstract terms, we present a deliberately simplified hypothetical example to illustrate the key concerns that motivated our approach to synthetic methods with compositional outcomes. Consider three countries A, B and C, where country A is affected by treatment at time T, while countries B and C were not treated and constitute the donor pool for creating synthetic controls for outcomes of interest. Suppose that we are interested in the treatment effects on party vote shares in country A, and all the voters in each country choose among three party families: X (e.g., left), Y (e.g., center-right), and Z (e.g., far right). Thus, the synthetic controls for vote shares of each party in country A are obtained by weighting the vote shares of the same party family in countries B and C.

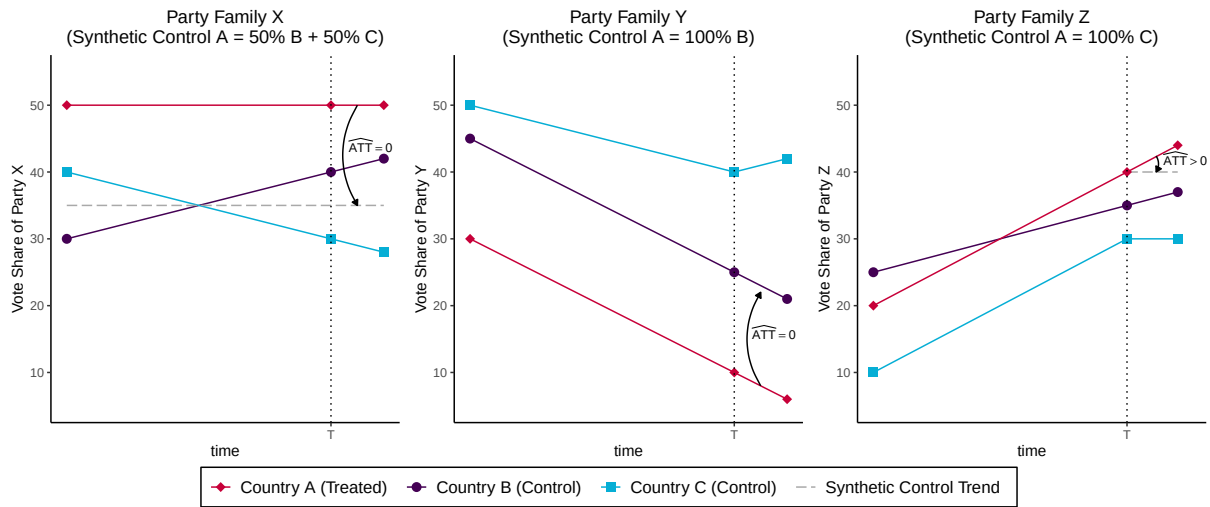


Figure 3.1. A hypothetical example illustrating the inconsistency in the weighting of control units (countries) for creating synthetic parallel trends for proportional outcomes (party vote shares).

In the approach to synthetic controls currently prevalent in political science, one would apply the method to each outcome variable separately (e.g., Cremaschi et al., 2024), often focusing on one party of interest (e.g., Becher et al., 2023; Hilbig and Riaz, 2024). Following this approach in a synthetic diff-in-diff framework, in our example, a researcher interested in the effects of the treatment would create a synthetic control for each party X, Y and Z separately, trying to approximate that party's pre-treatment trend in A by congenial party trends in B and C. For example, if A's pre-treatment trend is a bisector between the trends in B and C, then the optimal synthetic control for A comprises B and C half-and-half (Figure 3.1, left). However, this does not mean that synthetic controls for the vote shares of Y and Z will assign the same weights to countries B and C. The other two graphs in Figure 3.1 (center and right) show how this equivalence can easily be violated.

Thus, using the synthetic control methods on each party X, Y and Z separately can lead to effect estimates that do not add up to zero⁴ because each estimate is essentially obtained using a different control. Meanwhile, using the same synthetic control weights obtained from analyzing one outcome to estimate the effects on the related proportional

⁴Note that in Figure 3.1, vote shares of X, Y and Z for each country always add up to 100, as in the example these three party categories obtain all the votes in each country.

outcomes (e.g., vote shares for Y and Z using the weights from analyzing X) would make the effects compatible. However, such an approach could also come with downsides: the same weights that perfectly approximate the pre-treatment trend for X can imply quite imperfect synthetic controls for Y and Z. In what follows, we present this argument formally and propose an approach that takes into account proportional outcomes that form a whole.

3.3.2 Constant Control Comparison

We conceptualize the idea of constant control comparison for compositional data, focusing on a balanced panel with N observations at T time points set in the following way. For each unit $i \in \{1, \dots, N\}$ and each time point $t \in \{t, \dots, T\}$ we observe K proportions y_{itk} , with $k \in \{1, \dots, K\}$. We observe N_{tr} treated ($D_i = 1$) and N_{co} not treated control units ($D_i = 0$), where $i \in \{1, \dots, N_{co}, N_{co} + 1, \dots, N\}$. In this section, we limit the description to the case where exposure to treatment occurs at the last time point $T = T_0 + 1$.⁵

For each observation, we introduce potential outcomes under the control condition $y_{itk}(0)$ and under the treatment condition $y_{itk}(1)$. This formulation inherits the fundamental problem of causal inference that only one of the counterfactual outcomes for the different proportions can be observed (Holland, 1986). The key interest lies in the formulation of average effect estimates on each of the proportional outcomes. We are going to focus on the average treatment effects on the treated (ATT) for the different proportional outcomes: $\tau_k = \mathbb{E}[y_{iT_k}(1) \mid D_i = 1] - \mathbb{E}[y_{iT_k}(0) \mid D_i = 1]$.

Compositional data implies that the counterfactual and observed proportions are bound between 0 and 1 ($0 \leq y_{itk}(d) \leq 1$) and sum to 1 ($\sum_{k=1}^K y_{itk}(d) = 1$) for all observations i , time points t , and the two treatment states $d \in \{0, 1\}$. Lemma 1 describes that, under the proportional constraint, the sum of average treatment effects on the separate proportions should be zero.

Lemma 1 (Constraint on ATTs for proportions). *For compositional data where $0 \leq$*

⁵The results of this section can easily be generalized to multiple post-treatment periods.

$y_{itk}(d) \leq 1 \ \forall \ k$ and $\sum_{k=1}^K y_{itk} = 1$ it holds that

$$\sum_{k=1}^K \tau_k = \sum_{k=1}^K \mathbb{E}[y_{iT_k}(1) - y_{iT_k}(0) \mid D_i = 1] = 0, \quad (3.1)$$

where $y_{itk}(d)$ is the potential outcome under treatment status $d \in \{0, 1\}$.⁶

The results imply that if one proportion increases under treatment exposure, another has to decrease, leading to a sum constraint over the treatment effects of zero. Violations of the constraint on the ATTs lead to contradictory interpretations of the effects on different proportions and can make it difficult to interpret the results.

To see if the constraint on ATTs for proportions is met by causal inference methods in panel data settings, it is useful to consider the way how they estimate the unobserved counterfactual. For the ATTs, the outcomes for the treated units under treatment are observable and $\mathbb{E}[y_{iT_k}(1) \mid D_i = 1]$ can be estimated from the data in a consistent and unbiased way as $\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{iT_k}$. The key challenge lays in evaluating what the expected outcome for the treatment units would have been without exposure: $\mathbb{E}[y_{iT_k}(0) \mid D_i = 1]$.

Synthetic control methods (SCMs) and difference-in-difference (DID) methods can be understood as estimating the counterfactual as a weighting function of the control outcomes. To represent this idea, we work with a demeaned weighting estimator (Doudchenko & Imbens, 2016).

Definition 1. *The demeaned weighting estimator first computes the pre-treatment averages for the outcomes k , defined as $\bar{y}_{i,k} = \frac{1}{T_0} \sum_{t=1}^{T_0} y_{itk}$. It then calculates the demeaned outcomes as $\dot{y}_{itk} = y_{itk} - \bar{y}_{i,k}$. These pre-treatment averages and demeaned outcomes are used to create a linear combination of the pre-treatment averages and weighted post-treatment demeaned outcomes from the control units, which are then used to estimate the counterfactual outcomes for the treatment units:*

⁶Please refer to Appendix 3.A.2 for proofs.

$$\hat{y}_{iTk}(0) = \bar{y}_{i.k} + \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jTk} \quad (3.2)$$

for all $i \in \{N_{co} + 1, \dots, N\}$, where w_{jk} are a set of non-negative weights that sum to one: $\sum_{j=1}^{N_{co}} w_{jk} = 1$ for all k . The demeaned weighting estimator of the ATT is given by the difference between the average of the observed outcomes for the treated units and the average estimated counterfactual outcomes of the treated units:

$$\hat{\tau}_k = \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{iTk} \right) - \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left(\bar{y}_{i.k} + \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jTk} \right) \right). \quad (3.3)$$

Important in estimation is the way different methods choose the weights. DID methods usually assign equal weights to all control units, such that $w_{jk} = 1/N_{co}$. This means that for all units in the control group and all proportional outcomes, the control weights are the same: $w_{jk} = w_{jk'}$ for all $j \in \{1, \dots, N_{co}\}$ and $k, k' \in \{1, \dots, K\}$. However, SCMs estimate weights to match the pre-treatment outcomes, as formally discussed in Appendix 3.A.1. This leads to weights that are not necessarily equal for the control units, and neither for all proportions within a control unit. A separate application of SCMs to proportional outcomes can yield weights that can differ for the same unit for different proportional outcomes: there exist $k, k' \in \{1, \dots, K\}$ with $k \neq k'$ such that $w_{ik} \neq w_{ik'}$.

In contrast, we define the constant control comparison for compositional data in the following way.

Definition 2. *Constant control comparison imposes the same weights on control units for all proportional outcomes $w_{ik} = w_{ik'}$ for all $k, k' \in \{1, \dots, K\}$ with $k \neq k'$ in the estimation of the counterfactual expectation $\hat{y}_{iTk}(0)$ using the demeaned weighting estimator (see Definition 1).⁷*

This constant control comparison, in the form of the same weights on control units

⁷The constant control comparison enables using the same weights for all proportional outcomes of a unit, denoted as w_j in the demeaned weighting estimator, rather than specific weights for the proportional outcome w_{jk} .

for different proportional outcomes, ensures that the constraint on ATTs for proportional outcomes is met.

Theorem 1. *The constant control comparison assures that the constraint on ATTs for proportions holds (see Lemma 1). Given Definition 2 it holds that the sum of the demeaned weighting estimates of the ATT (see Definition 1) are equal to zero:*

$$\sum_{k=1}^K \hat{\tau}_k = \sum_{k=1}^K \left(\left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{iTk} \right) - \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left(\bar{y}_{i,k} + \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk} \right) \right) \right) = 0 \quad (3.4)$$

If the constant control comparison is violated, the constraint on ATTs for proportions does not necessarily hold. For an illustrative example of the violation of the constant control comparison with weights that differ for proportions, see Appendix 3.A.2.

This leads to the central insight that SCMs aiming to satisfy the constraint on the ATTs for compositional data should estimate identical weights for all proportions to ensure the constant control comparison holds. In contrast, the common practice of applying SCMs separately to proportional outcomes can violate this constraint on the ATTs.

3.3.3 Common Synthetic Control Weights Across Outcomes

Following Definition 2, any weights w_{ik} for control units that are common among the proportional outcomes $\forall k \in \{1, \dots, K\}$ of the units satisfy the constant control comparison. However, the requirement to hold the weights constant alone does not yet imply which weights should be used.

As discussed previously, in DID methods, common weights for different outcomes are implicitly imposed since all control observations receive the same weight. Researchers sometimes also implicitly select weights by pre-selecting a control group before estimating the effect. For example, Cremaschi et al. (2025) estimate ATTs of plant disease on party

vote shares in the Italian region of Apulia, using uninfected Apulian municipalities as a DID control group across different parties. The choice of the control group for compositional outcomes can be data-driven: Selb and Munzert (2018) use either all German counties not exposed to Hitler’s speeches or only their subsample matched by covariates to estimate the speech on the votes for Nazis and Communists. Regardless of how the control is chosen, the researcher can hold the weights constant using the same control group for the effects on different proportions.

However, SCMs for univariate outcomes cannot guarantee common control weights across different outcomes⁸. With SCMs, calculating these weights is a part of each estimation. Thus, starting with the same sample of treated and untreated (donor pool) units for different outcomes, one ends up with weights that fit this outcome best.

To ensure common weights across different outcomes, the synthetic control approaches must be modified. One solution previously adopted by researchers is calculating the weights for just one outcome and using them across outcomes. Indeed, after calculating synthetic control weights based on the regional GDP and its predictors, Pinotti (2015) uses the same weights to estimate the effects of organized crime on the GDP and the homicide rate. Such “incomplete” application of SCMs to all but one outcome resembles matching on covariates with subsequent DID application across outcomes, similar to Selb and Munzert (2018)’s approach to parties in interwar Germany. However, deriving synthetic control weights from only one outcome variable risks over-fitting its pre-treatment trends (Abadie, 2021) and neglects the core purpose of SCMs for other outcomes.

We propose that for compositional data, synthetic control weights should maximize the goodness of pre-treatment fit for all proportional outcomes taken together. Indeed, in a DID setup, one would require the parallel trends assumption to hold for each proportional outcome for valid inference. As violations of parallel trends in the pre-intervention (multivariate) outcomes invalidate the use of DID and make the case for synthetic con-

⁸Unless there are only two proportional outcomes that sum to 1, so that knowing one proportion always means knowing both. In this case, the constant control comparison will hold.

trols, the latter should aim to maximize the pre-treatment fit for all proportional outcomes – simultaneously, i.e. using one common set of weights.

In a recent paper, Sun et al. (2023) demonstrate that when applying SCMs to multiple outcomes, estimating a common set of weights across outcomes can indeed be beneficial. They show that, under a low-rank factor model, a common set of weights leads to lower bias bounds compared to separate weights. Furthermore, they find that averaging these weights provides additional gains as the number of outcomes increases. These insights are also relevant to our work, as they show that a common set of weights not only addresses the inherent constraint of treatment effects in compositional data but also may lead to more reliable inferences. The next section formalizes our approach to calculating synthetic control weights.

3.4 Synthetic Control Methods for Proportions

3.4.1 Estimation with Constant Control Weights

Based on the insights and the motivating example, we formulate a synthetic control method for proportions that follows the constant control comparison. The formulation of it closely follows Arkhangelsky et al. (2021), except for the fact that we calculate the same set of pre-treatment weights for all proportions.

We again consider a balanced panel setup with N units and T time periods. We denote units by $i \in \{1, \dots, N\}$ and periods with $t \in \{1, \dots, T\}$. The outcomes of interest are K proportions, where each proportion value is denoted by $k \in \{1, \dots, K\}$. The outcome for unit i in period t is a multivariate vector of $Y_{it} = [y_{it1}, \dots, y_{itK}]$. For compositional data of proportion, it holds that $y_{itk} \in [0, 1]$ for all k and $\sum_{k=1}^K y_{itk} = 1$, forming a $(K - 1)$ -dimensional simplex. We also have a binary treatment $D_{it} \in \{0, 1\}$. We assume that there is a share of control units that are never exposed to treatment N_{co} , while there are units N_{tr} that are treated after time T_{pre} . For the analysis, we arrange the i units such that first all control units appear and then all treatment units

$i \in \{1, \dots, N_{co}, N_{co} + 1, \dots, N\}$.

Synthetic methods calculate a set of non-negative weights $\hat{\omega}_{ik}$ that align pre-exposure trends in the control outcomes with the treatment units, such that $\sum_{i=1}^{N_{co}} \hat{\omega}_{ik} y_{itk} \approx \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk}$ for all $t \in \{1, \dots, T_{pre}\}$. As Arkhangelsky et al. (2021) describe, the idea of weighting control units to align with the pre-exposure trends of the treated units can be applied to both the classic synthetic control method, where weights are calculated such that the pre-trends are matched directly, and synthetic DID methods, where weights are calculated such with a parallel shift and additional time weights. For more details on the models, see Appendix 3.A.1.

In our application of synthetic control methods to proportions, we calculate the same set of individual weights for all proportional outcomes $\hat{\omega}_i$, and do the same for time weights $\hat{\lambda}_t$ in the case of synthetic DID. The estimation of the weights has a minimization problem of the following form.

$$\sum_{i=1}^{N_{co}} \hat{\omega}_i y_{itk} \approx \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk} \quad \forall t \in \{1, \dots, T_{pre}\} \wedge k \in \{1, \dots, K\} \quad (3.5)$$

This leads to a minimization problem, where we minimize squared deviations of overall proportions with one set of non-negative weights.

$$(\hat{\omega}_0, \hat{\omega}) = \arg \min_{\hat{\omega} \in \mathbb{R}, \hat{\omega} \in \Omega} \ell_{uni}(\omega_0, \omega) \text{ where} \quad (3.6)$$

$$\ell_{uni}(\omega_0, \omega) = \sum_{k=1}^K \left(\sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i y_{itk} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk} \right)^2 \right) + \xi^2 T_{pre} \|\omega\|_2^2,$$

$$\Omega = \left\{ \omega \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_i = 1, \omega_i = \frac{1}{N_{tr}} \quad \forall i \in \{N_{co} + 1, \dots, N\} \right\}$$

Important here is the summation over the pre-treatment periods and the different proportions K with respect to the same set of weights for all proportions.

The formulation also includes a regularization parameter that penalizes weights that

are too large using a L2 norm. The use of a regularization parameter, ξ , helps ensure the uniqueness and stability of the estimated weights by controlling the dispersion, which leads to more robust estimations (see also Doudchenko & Imbens, 2016).⁹

The time weights λ pose a similar adapted minimization problem:

$$(\hat{\lambda}_0, \hat{\lambda}) = \arg \min_{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda} \ell_{time}(\lambda_0, \lambda) \text{ where} \quad (3.7)$$

$$\begin{aligned} \ell_{uni}(\lambda_0, \lambda_i) &= \sum_{k=1}^K \left(\sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t y_{itk} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T y_{itk} \right)^2 \right), \\ \Lambda &= \left\{ \lambda \in \mathbb{R}_+^T : \sum_{i=1}^{T_{pre}} \lambda_t = 1, \lambda_t = \frac{1}{T_{post}} \forall i \in \{T_{pre} + 1, \dots, T\} \right\} \end{aligned} \quad (3.8)$$

The minimization problem for the time weights does not include a penalization parameter, as in (Arkhangelsky et al., 2021).

The weights for unit and time can then be used in the two-way fixed effects model to use the synthetic DID estimator. The two-way fixed effect regression is of form:

$$(\hat{\tau}_k^{sdid}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \hat{\omega}_i \hat{\lambda}_t \right\}, \quad (3.9)$$

where α_{ik} are unit fixed effects for each proportion, and β_{tk} are time fixed effects for all proportions.

Algorithm 1 describes the overall procedure for the Synthetic Difference in Differences (SDID) Estimator with common weights. We first calculate the regularization

⁹The regularization parameter is pre-specified in the same way as in standard SCMs and described in the Appendix 3.A.1

Algorithm 1. Compute the Synthetic Difference in Differences (SDID) Estimator

- 1: Compute regularization parameter ξ
- 2: Compute unit weights $\hat{\omega}$
- 3: Compute time weights $\hat{\lambda}$
- 4: **for** each proportion $k \in \{1, \dots, K\}$ **do**
- 5: Compute the SDID estimator via the weighted DID regression:

$$(\hat{\tau}_k^{sdid}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \hat{\omega}_i \hat{\lambda}_t \right\} \quad (3.10)$$

- 6: **end for**
-

parameter and then compute the unit weights followed by the time weights. Afterward, separately estimate the treatment effects using the weighted two-way regression for each proportion. The result is a set of treatment estimates on the proportions.¹⁰ We implement both SCMs for compositional proportional outcomes in an R-package (`propsdid`).¹¹

3.4.2 Inference

To estimate the uncertainty around our treatment effect estimates SCMs models for proportions, we rely on the same approaches to variance estimation as in Arkhangelsky et al. (2021). These methods offer valid inference under large-sample conditions and are also well suited to our SCMs framework with constant control weights.

One widely used method is clustered bootstrap (Efron, 1979), which accounts for the correlation between units by resampling units with replacement. For each bootstrap sample, we recompute the SCMs estimator and estimate the variance of the treatment effect. Although computationally intensive, especially in large panels, this approach has been shown to perform robustly in a variety of settings (Bertrand et al., 2004).

¹⁰A similar procedure can be used to estimate classic synthetic control method models, usually applied in context with one treatment unit (Abadie et al., 2010, 2015). For this, we skip the time weight estimation, excluding the time weights from the regression model. We further exclude the constant in the calculation of the unit weights, which does not allow for an intercept shift in pre-treatment trends. A description of the algorithm is provided in Appendix 3.A.1.

¹¹The package extends `synthdid` (Hirshberg, 2022) to include functionality for the multivariate outcomes with constant control comparisons.

The application of jackknife methods (Miller, 1974), as discussed by Arkhangelsky et al. (2021), reduces the computational cost by requiring that the full SDID algorithm be run only once. It systematically omits individual units to provide a conservative variance estimate. This method works particularly well when control and treated units are sufficiently similar, allowing the time weights to generalize effectively between them. Under specific conditions, the jackknife can yield exact confidence intervals, making it a reliable choice for inference.

Placebo inference, meanwhile, is particularly valuable when there are very few treated units, such as when there is only one (Abadie et al., 2010). As highlighted in Arkhangelsky et al. (2021), this method involves applying placebo treatments to control units to estimate the level of noise, similar to the approach of Conley and Taber (2011). Placebo inference assumes homoskedasticity across units, making it well-suited for scenarios with sparse treatment exposure.

Each of these methods strikes a balance between computational complexity and robustness, depending on the structure of the dataset. For large panels, bootstrap is often favored for its reliability, whereas the jackknife and placebo methods provide useful alternatives when computational efficiency or small-sample concerns are more pressing. In our application, we rely on bootstrap method.

3.5 Simulation Study

In this section, we evaluate the performance of the synthetic control method estimator for proportions. We use a dynamic latent state model to establish the data-generating process (DGP) for our simulation study, which produces balanced panel data with T time points, N units, and K proportions. Units are assigned to treatment probabilistically based on latent factors, with treatment starting at time T_1 for N_1 units. The latent unit factors evolve over time, and outcomes are simulated using a multinomial link to ensure proportional outcomes. This setup enables us to model violations of the parallel trends assumption while accounting for treatment selection based on either initial local level

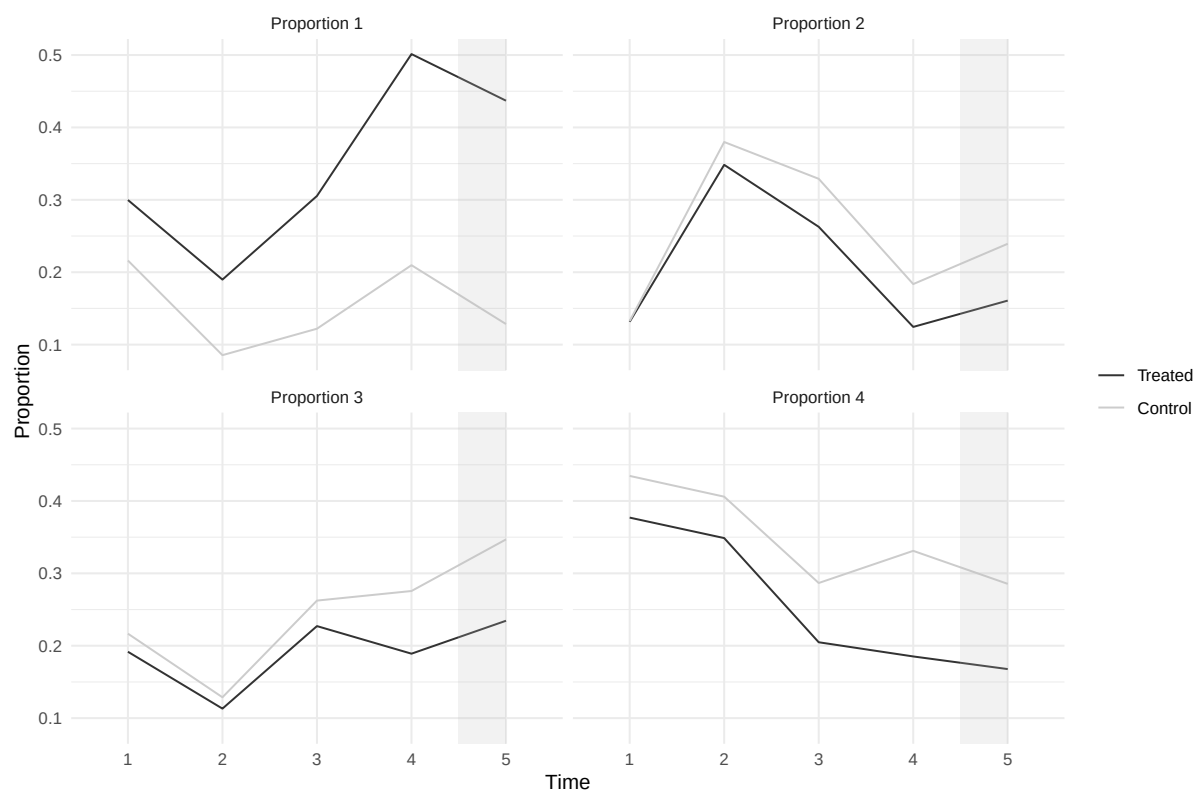


Figure 3.2. Example Simulation data from the data generating process (DGP). We analyze a situation with 4 proportions, and a latent treatment vector of $[0.5, 0, 0, 0]$. The example is for 400 units (of which 200 are treated) and 5 time points, where in the last time points the treatment is introduced. The first proportion exhibits a local trend, and selection is based on the local trend.

differences or differences in local trends. In particular, local trends that, on average, differ between control and treatment units violate the assumption of parallel trends, requiring synthetic methods to adjust for this violation. Additionally, selection based on local levels can also violate the parallel trends assumption due to the non-linear relationship between these levels and the proportional outcomes. Detailed steps and parameter specifications are provided in the Appendix [3.A.4.1](#).

For analysis, we analyze a scenario with four proportions and a latent treatment effect on the first proportion. Figure [3.2](#) illustrates the average outcomes for treated and control units over time from an example data set with 400 units, of which 200 are treated at the final time point of the five observed. The effect in this figure is not as clear but indicates an increase in proportion one. It also shows that treatment selection

based on local trends leads to a violation of the parallel trends assumption and that with compositional data, the effects we induce for proportion one affect the processes for the other proportional outcomes.

For the general analysis, we set up the data generation parameters in the following way. We analyze scenarios with 5 and 10 time points, where the treatment occurs either in the final period or in the last two periods. We consider 50, 200, and 1000 units, with treatment proportions of 10% and 50%. Treatment selection is done in two ways: once based on local levels and once based on local trends.¹² We draw 1000 independent samples for each of 48 combinations and estimate a difference-in-difference (DID) model using two-way fixed effects, a synthetic control (SC) and a synthetic difference-in-differences (SDID) model separately for all outcomes as well as our proposed synthetic control (PropSC) difference-in-differences (PropSDID) models for proportions. We compare the estimates based on standard deviation (SD), root-mean-square error (RMSE), and absolute-valued bias (abs. bias). We also calculate the absolute deviation of the sum of the treatment effect estimates on the separate proportions (Sum const.), which under the constant control comparisons should be zero¹³.

Table 3.1 describes the results. We aggregate the results for all combinations of parameter setups, but differentiate between the treatment selection based on local trends and local levels, aggregating the metrics over all K proportions. First, the table shows that the constant comparison group indeed leads to the sum constraint holding. For difference-in-differences (DID), and proportional synthetic control (PropSC) and proportional synthetic difference-in-differences (PropSDID), the sum of the treatment effects on the different proportions is always zero, while for the separate application of synthetic difference-in-differences (SDID) and synthetic control, there are deviations on average. For the simulations with treatment selection based on local levels, the average deviation is 1.4 percentage points for SDID and 5.8 percentage points for SC, and for the selection based on local trends, it is 2.4 (SDID) and 5.0 (SC) percentage points. Given that the

¹²We further fix the other parameters of the DGP as described in Appendix 3.A.4.2.

¹³We calculate the true ATTs in the sample by comparing the counterfactuals for treated units in treatment time points.

Method	SD	RMSE	Abs. bias.	Sum const.
Treatment selection: local levels				
DID	7.2	5.5	4.1	0
SC	5.8	5.4	3.9	5.8
PropSC	5.4	4.9	3.6	0
SDID	4.5	3.6	2.5	1.4
PropSDID	4.4	3.6	2.5	0
Treatment selection: local trends				
DID	7.2	7.2	5.5	0
SC	5.8	5.1	3.6	5
PropSC	5.4	4.8	3.3	0
SDID	4.5	4.0	2.8	2.4
PropSDID	4.4	3.8	2.6	0

Table 3.1. Monte Carlo evaluation of difference-in-differences (DID), synthetic difference-in-differences (SDID), proportional synthetic difference-in-differences (PropSDID), synthetic control (SC), and proportional synthetic control (PropSC). For treatment selection on local levels and local trends, we compute the standard deviation (SD), root mean square error (RMSE), and absolute bias (Abs. Bias), along with the sum constraint (Sum Const.). All values are scaled by 100. The results are based on 96,000 Monte Carlo samples, with a Monte Carlo standard error below 0.05 for all metrics.

average treatment effect on the treated (ATT) in our simulation is 3.6 percentage points, this is a substantial deviation, which can make it difficult to interpret the results from the separate application of DID accurately.

Second, SCMs are superior to DID, which has a considerable bias of 4.1 percentage points when treatment selection is based on local levels, and an even higher bias of 5.6 percentage points when based on local trends. This bias is also reflected in high root-mean-square error (RMSE).

Synthetic control methods are generally preferable under our data-generating process, as expected given the violation of parallel trends. The absolute error remains below 4 percentage points for both SC and PropSC and even falls below 3 points for SDID and PropSDID, underscoring the superior performance of the synthetic DID model compared to classic synthetic control methods in our simulations. However, we observe only a marginally smaller RMSE for the proportional versions of both the synthetic difference-

in-differences and the synthetic control methods. For local levels, the RMSE for SDID methods does not show a significant improvement with proportional adjustment. However, in local trends, the RMSE for SDID is 0.04, slightly higher than the 0.038 for PropSDID, and 0.51 for SC versus 0.48 for PropSC. These differences are relatively minor and may be negligible in practical applications.¹⁴

Overall, the simulations confirm the central role the constant control comparison plays in ensuring the sum constraint of the ATT on proportional outcomes. They also highlight that our proportional synthetic control estimation, which ensures that the control comparison holds, is at least on par with the separate application of the method in our simulations.

3.6 Empirical Applications

3.6.1 Application 1: Climate Policy and Voting Outcomes

In this section, we apply our proportional synthetic control methods framework to reevaluate the findings presented by Bolet et al. (2024) in the *American Political Science Review*. We use this example for three purposes. First, we demonstrate the importance of considering multivariate outcomes instead of focusing on one proportion, as previously discussed by political scientists analyzing compositional data (Katz & King, 1999; Lipsmeyer et al., 2019; Philips et al., 2015, 2016). Second, we demonstrate the imbalance of the effect estimates obtained from applying SCMs to compositional variables separately. Third, we show how this issue can be resolved by using the proportional method described in the current paper.

Bolet et al. (2024) examine the electoral effects of a compensatory “Just Transition Agreement” negotiated by the Spanish government in coal-mining areas transitioning away from environmentally harmful production. Using electoral data from 2008 to 2019 and focusing on the 2019 electoral result of the governing Spanish Socialist Workers’ Party

¹⁴We also find that the bias decreases with a larger sample size, particularly in the number of control units (See Appendix 3.A.4.3).

	PSOE	PP	PODEMOS	Citizens	VOX	Others
DID Estimates						
2019 $\times$ Coalmines	1.825* (0.771)	1.731 (0.940)	0.447 (0.398)	0.669 (0.814)	-4.593*** (0.482)	-0.071 (0.531)
SDID Estimates						
2019 $\times$ Coalmines	1.82** (0.65)	1.52* (0.76)	0.39 (0.34)	0.69 (0.65)	-3.89*** (0.47)	-0.04 (0.29)
PropSDID Estimates						
2019 $\times$ Coalmines	1.33* (0.69)	1.08 (0.82)	0.26 (0.36)	0.77 (0.69)	-3.41*** (0.53)	-0.03 (0.26)
Num. Obs.	2455					

Table 3.2. The estimates of the effects of the Just Transition Agreement on party vote shares in 2019 Spanish elections, using DID, SDID, and proportional SDID approaches.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(PSOE) as the outcome of interest, the study applies a DID framework to compare changes in voting between mining and non-mining municipalities in three Spanish provinces. With a few DID specifications arriving at similar point estimates of plus 1.8 percentage point effect on PSOE’s vote share, the authors interpret that as the electoral reward from the voters to the incumbent party for the Just Transition Agreement.

To reassess these findings, we first replicate the analysis by Bolet et al. (2024) for a complete set of party vote shares that sum up to 100: PSOE (the aforementioned center-left governing party), PP (People’s Party, center-right opposition), PODEMOS (left opposition), Citizens (right-liberal opposition), VOX (far-right opposition) and others (the residual category including all remaining small parties). For this, we downloaded the electoral results for 2008-2019 national elections from the Spanish Electoral Archive (Pérez et al., 2021). Our replication confirms the reproducibility of the findings by Bolet et al. (2024) for PSOE vote shares.¹⁵ Now, we estimate the effect of the Just Transition

¹⁵We base the DID replication on the way the models are specified in the replication package. While the article text states that model 1 includes municipal fixed effects, the code is missing them, thus we reproduce the code and not the paper description. The numerical results of Bolet et al. (2024) can only be reproduced, provided we also deliberately introduce a data error from the replication package into our independently cleaned electoral data, understating PSOE’s result in one of the municipalities in 2008 by about 15 percentage points. Without this error, all point estimates from Bolet et al. (2024) become slightly smaller but remain close to the original results. Hereinafter, we use the data without that error.

Agreement on a complete set of parties using the DID specification from the original study that also includes province and year fixed effects as well as covariates such as the log population, unemployment growth, the shares of immigrants, people with primary education and men over 50 in the population, all provided by the authors in their replication package. The results of this estimation, presented in Table 3.2, paint an interesting picture: most notably, the radical right VOX has suffered a large negative effect of minus 4.6 percentage points. The votes lost by VOX did not only benefit PSOE by about 1.8 percentage points, as was already established by Bolet et al. (2024). The results also indicate an almost equally large positive effect, significant at 0.1 level, on the vote share of the center-right opposition PP. Consistent with the properties of DID we discussed above, the sum of the effects for an exhaustive set of compositional outcomes is equal to zero, i.e. the effects on different parties are consistent with each other and some parties' losses are translatable to other parties' gains.

Now, we apply the synthetic difference-in-differences method to the same case. As SDID is a generalization of DID methods that relaxes the assumptions about the raw data and typically outperforms DID (Arkhangelsky et al., 2021), SDID is generally applicable to any DID setting. The pre-treatment trends in the voting outcomes between treated and control units are virtually never perfectly parallel, nor are they in our case, as can be seen in Figure 2 in Bolet et al. (2024). Thus, applying SDID can ensure the best pre-treatment fit in vote shares of the synthetic control municipalities and the coalmine municipalities for each party. The results of this estimation, presented in Table 3.2, paint a picture similar to the DID results in terms of size and sign of the coefficients. The 1.8 p.p. effects on PSOE is very close to the DID estimate, while PP has gained 1.5 p.p. and VOX has lost 3.9 p.p. from the treatment, according to the SDID estimates. However, as anticipated in Section 3.3, these effect estimates should be interpreted carefully as they are not consistent with each other: instead of summing up to zero, these effects sum up

Furthermore, it appears that the authors defined PSOE vote share as a ratio of PSOE votes to the number of all ballots cast, including blank and null (invalid) votes. This differs from the way vote shares are commonly calculated by political scientists and electoral statistics, dividing the votes by party only by valid votes (excluding null ballots). We use that latter approach to ensure a more intuitive interpretation of our estimates and their direct applicability to the official electoral results calculated by the Spanish authorities.

to $1.82 + 1.52 + 0.39 + 0.69 - 3.89 - 0.04 = 0.89$, which contradicts to the nature of proportions in compositional data.

To resolve this issue, we now estimate the same effects with our novel synthetic control methods tailored for multivariate proportional outcomes, using the `propsdid` package accompanying this paper. Thus, we ensure a constant synthetic control group for the coalmine municipalities, aligning the effects with the essential sum constraint where any increase in vote shares for one party necessitates a decrease in another. The results of this estimation, presented in Table 3.2, are generally more conservative than any of the previous estimates, pointing that those could overestimate the effects. According to the proportional SDID results, the overall negative effects on the far right VOX was about -3.4 percentage points. Though the positive effects on the PP is not significant on the 5%-level, the point estimates for the largest parties PSOE and PP remain positive and larger than for the remaining parties, suggesting that PSOE and PP were the main beneficiaries of the VOX losses.

Our analysis demonstrates the applicability of proportional synthetic control methods for causal inference with panel data on compositional outcomes. Moreover, in that case, the estimates provided by proportional SDID are more reliable than both DID and SDID, as our approach maximizes the fit in parallel pre-treatment trends while keeping the effects balanced with each other. The evidence obtained from this application complements and challenges the initial conclusions drawn by Bolet et al. (2024), showing that the Just Transition Agreement did not just draw voters to reward to governing party. Rather, it had a deradicalizing effect, making the citizens vote less for the far right and more for the large mainstream parties both in government and opposition. Besides, applying the synthetic DID estimator shows that those effects appear to be smaller than their estimates obtained by the DID methods.

3.6.2 Application 2: Electoral Effects of Anti-LGBTQ Policies

In this section, we apply the novel synthetic control methods tailored for multivariate proportional outcomes to evaluate the impact of anti-LGBTQ resolutions on electoral

outcomes in Poland, replicating and extending the analysis from Chapter 2. In the second chapter, we examined the electoral consequences of anti-LGBTQ resolutions adopted by Polish municipalities before the 2019 parliamentary election. Using a synthetic difference-in-differences (SDID) approach, we found that municipalities adopting these resolutions experienced a significant decrease in overall turnout and the vote shares for some opposition parties, while the ruling far-right Law and Justice Party (PiS) gained a vote share increase. For more detailed results, please refer to Chapter 2.

As the voting shares from Polish elections are compositional data, applying a synthetic difference-in-differences (SDID) approach has expectedly caused issues described in this chapter's Section 3.3. To improve the robustness and consistency of our findings, we now apply the synthetic control methods designed for proportional outcomes to the same data. This approach ensures that the estimated effects on different party vote shares adhere to the inherent sum constraint in compositional data.

Table 3.3 compares the results obtained using the traditional SDID method with those derived from our new proportional synthetic control method. The estimates are based on the full sample of Polish municipalities.

Comparing the results from Chapter 2 with the results from SDID for proportions, one can observe the following:

- **Law and Justice (PiS):** Both methods indicate an increase in PiS vote share among all ballots cast, with the PropSDID method showing a slightly larger effect. However, when considering the vote shares among eligible voters, the PropSDID method suggests a negative effect, highlighting the nuances captured by the new approach.
- **Civic Coalition (KO):** The traditional SDID method indicated a positive effect on KO's vote share among all ballots cast and no losses in votes, which was difficult to interpret. The PropSDID method shows an insignificant effect on the vote share and losses in turnout, which is more consistent with the effects on other mainstream opposition parties.

Party	SDID Vote Shares	PropSDID Vote Shares
Law and Justice (PiS)	0.82 (0.29)	0.93 (0.33)
Civic Coalition (KO)	0.56 (0.18)	-0.22 (0.16)
Polish Coalition (KP)	-1.98 (0.26)	-1.00 (0.33)
The Left (Lewica)	-1.26 (0.13)	-0.35 (0.12)
Other parties	-	0.63 (0.18)
Party	SDID Vote Shares (Eligible Voters)	PropSDID Vote Shares (Eligible Voters)
Law and Justice (PiS)	0.41 (0.18)	-0.28 (0.21)
Civic Coalition (KO)	-0.25 (0.09)	-0.91 (0.11)
Polish Coalition (KP)	-1.07 (0.12)	-0.75 (0.15)
The Left (Lewica)	-1.04 (0.07)	-0.42 (0.06)
Other parties	-	4.78 (0.29)
Turnout	-1.69 (0.11)	-2.42 (0.15)

Note: Jackknife standard errors ($\alpha = 5\%$) in parentheses.

Table 3.3. SDID vs PropSDID estimation results: Poland 2019

- **Polish Coalition (KP) and The Left (Lewica):** Both methods show significant negative effects on vote shares for these parties, consistent across the two approaches.
- **Other Parties:** The PropSDID method reveals a notable positive effect on the vote shares for other parties among eligible voters, not captured by the traditional SDID method.

Applying the synthetic control methods for proportional outcomes enhances the reliability and interpretability of our findings. Importantly, our main effect of interest on the vote share of the far-right PiS party remained very close across the two methods. However, we also received an even stronger indication of demobilizing (negative) effect on the overall turnout, and it became clear that none of the opposition parties who did not support the resolutions received an electoral gain. If anyone did, this was actually “other parties”, the label which mostly stands for the extreme right Confederation party, which also promotes homophobia. The mainstream opposition parties have lost from anti-LGBTQ resolutions, and the losses appeared more evenly spread across those parties than in Chapter 2.

3.7 Discussion

In this paper, we extend the synthetic control methods framework to handle compositional data as outcomes. Our proposed approach, which ensures constant control comparison across different proportions, adheres to the essential sum constraint inherent in compositional data. This enhances the validity and reliability of the treatment effect estimates and provides a more coherent interpretation of the results.

While existing methods, such as DID and synthetic controls with separate outcome modeling, often fall short when applied to compositional data, our approach fulfills the critical need for a valid methodology. Even for univariate outcomes, the parallel trends assumption required for DID is often implausible (Chiu et al., 2024; Hassell & Holbein, 2024), which makes it even less likely to hold for each proportion under study in a compositional context. Although synthetic control methods are a promising alternative, their separate application to each outcome violates the constant control comparison, as we demonstrated. Our synthetic control methods for proportions resolve this issue, providing a robust framework that addresses the unique challenges presented by compositional outcomes.

Our study builds on a long methodological tradition to take compositional outcomes seriously (Aitchison, 1982; Arnold et al., 2020; Cohen & Hanretty, 2024; Jackson, 2002; Katz & King, 1999; Lipsmeyer et al., 2019; Philips et al., 2016; Tomz et al., 2002). The formulated framework demonstrates a way to include compositional outcomes in causal inference with time-series cross-sectional data, central to many recent methodological advances (Arkhangelsky & Imbens, 2024; Blackwell & Glynn, 2018; Imai & Kim, 2019; Imai et al., 2023; Liu et al., 2024; Roth et al., 2023; Xu, 2024). Whenever proportions are studied, our approach can serve as a building block for a whole range of synthetic control methods - from the standard SC (Abadie et al., 2010, 2015) to generalized (Xu, 2017), penalized (Abadie & L'Hour, 2021) and augmented (Ben-Michael et al., 2021) synthetic controls as well as SDID (Arkhangelsky et al., 2021).

Our approach inherits the data requirements and assumptions of the synthetic control method implemented by the researcher. However, compared to separate outcome modeling, our approach to proportions potentially reduces the necessity for extensive pre-treatment data for each outcome (Abadie, 2021). Relying on the deterministic linear relationship among the proportions and the constant control comparison, one can potentially construct synthetic controls even for outcomes with limited pre-exposure data, such as a vote share for a recently formed political party. For instance, while new parties may have received 0% of the vote in past elections in all districts, our approach allows for the calculation of synthetic control weights based on the voting results of other parties. This capability is akin to using strong predictors of post-intervention values to mitigate the challenges posed by the lack of pre-exposure periods (Abadie, 2021). Future research should evaluate these potential benefits of further.

The implications of our research extend beyond synthetic control methods. The constant control comparison principle we outlined can be vital for other causal inference techniques that model or impute counterfactual outcomes, such as matching (Imai et al., 2023; Kellogg et al., 2021), trajectory balancing (Hazlett & Xu, 2018) or imputation-based methods (Athey et al., 2021; Borusyak et al., 2024; Liu et al., 2024). For instance, in matching techniques where researchers study the effect of covariates on voting decisions, applying separate matching weights for binary representations of these decisions can lead to estimates that violate the sum constraint, rendering the results difficult to interpret. We highlight the importance of adapting causal inference approaches for producing valid and meaningful results with compositional data.

References

- Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2), 391–425. <https://doi.org/10.1257/jel.20191450>
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59(2), 495–510. <https://doi.org/10.1111/ajps.12116>
- Abadie, A., & L’Hour, J. (2021). A Penalized Synthetic Control Estimator for Disaggregated Data. *Journal of the American Statistical Association*, 116(536), 1817–1834. <https://doi.org/10.1080/01621459.2021.1971535>
- Aitchison, J. (1982). The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society. Series B (methodological)*, 44(2), 139–177. <https://doi.org/10.1111/j.2517-6161.1982.tb01195.x>
- Alcocer, M. (2024). Increasing Intergovernmental Coordination to Fight Crime: Evidence from Mexico. *Political Science Research and Methods*, 1–10. <https://doi.org/10.1017/psrm.2024.31>
- Alrababa’h, A., Marble, W., Mousa, S., & Siegel, A. A. (2021). Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes. *American Political Science Review*, 115(4), 1111–1128. <https://doi.org/10.1017/S0003055421000423>
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, 111(12), 4088–4118. <https://doi.org/10.1257/aer.20190159>
- Arkhangelsky, D., & Imbens, G. (2024). Causal models for longitudinal and panel data: A survey. *The Econometrics Journal*, 27(3), C1–C61. <https://doi.org/10.1093/ectj/utae014>
- Arnold, K. F., Berrie, L., Tennant, P. W. G., & Gilthorpe, M. S. (2020). A Causal Inference Perspective on the Analysis of Compositional Data. *International Journal of Epidemiology*, 49(4), 1307–1313. <https://doi.org/10.1093/ije/dyaa021>
- Athey, S., Bayati, M., Doudchenko, N., Imbens, G. W., & Khosravi, K. (2021). Matrix Completion Methods for Causal Panel Data Models. *Journal of the American Sta-*

- tistical Association*, 116(536), 1716–1730. <https://doi.org/10.1080/01621459.2021.1891924>
- Athey, S., & Imbens, G. W. (2017). The econometrics of randomized experiments. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Economic Field Experiments* (pp. 73–140). North-Holland. <https://doi.org/10.1016/bs.hefe.2016.10.003>
- Becher, M., González, I. M., & Stegmueller, D. (2023). Proportional Representation and Right-Wing Populism: Evidence from Electoral System Change in Europe. *British Journal of Political Science*, 53(1), 261–268. <https://doi.org/10.1017/S0007123421000703>
- Ben-Michael, E., Feller, A., Hirshberg, D. A., & Zubizarreta, J. R. (2021). *The Balancing Act in Causal Inference*. arXiv: 2110.14831. <https://doi.org/10.48550/arXiv.2110.14831>
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1), 249–275. <https://doi.org/10.1162/003355304772839588>
- Blackwell, M., & Glynn, A. N. (2018). How to Make Causal Inferences with Time-Series Cross-Sectional Data under Selection on Observables. *American Political Science Review*, 112(4), 1067–1082. <https://doi.org/10.1017/S0003055418000357>
- Bolet, D., Green, F., & González-Eguino, M. (2024). How to Get Coal Country to Vote for Climate Policy: The Effect of a “Just Transition Agreement” on Spanish Election Results. *American Political Science Review*, 118(3), 1344–1359. <https://doi.org/10.1017/S0003055423001235>
- Borusyak, K., Hull, P., & Jaravel, X. (2024). Design-based identification with formula instruments: A review. *Econometrics Journal*, utae003. <https://doi.org/10.1093/ectj/utae003>
- Chiu, A., Lan, X., Liu, Z., & Xu, Y. (2024). *What To Do (and Not to Do) with Causal Panel Analysis under Parallel Trends: Lessons from A Large Reanalysis Study*. <https://doi.org/10.2139/ssrn.4490035>
- Cohen, D., & Hanretty, C. (2024). Simulating Party Shares. *Political Analysis*, 32(1), 140–147. <https://doi.org/10.1017/pan.2023.13>
- Conley, T. G., & Taber, C. R. (2011). Inference with “Difference in Differences” with a Small Number of Policy Changes. *Review of Economics and Statistics*, 93(1), 113–125. https://doi.org/10.1162/REST_a_00049
- Cremaschi, S., Bariletto, N., & De Vries, C. E. (2025). Without Roots: The Political Consequences of Collective Economic Shocks. *American Political Science Review*, 1–20. <https://doi.org/10.1017/S0003055425000073>
- Cremaschi, S., Rettl, P., Cappelluti, M., & De Vries, C. E. (2024). Geographies of discontent: Public service deprivation and the rise of the far right in Italy. *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12936>
- Currie, J., Kleven, H., & Zwiers, E. (2020). Technology and Big Data Are Changing Economics: Mining Text to Track Methods. *AEA Papers and Proceedings*, 110, 42–48. <https://doi.org/10.1257/pandp.20201058>

- de Kadt, D., Johnson-Kanu, A., & Sands, M. L. (2024). State Violence, Party Formation, and Electoral Accountability: The Political Legacy of the Marikana Massacre. *American Political Science Review*, 118(2), 563–583. <https://doi.org/10.1017/S0003055423000448>
- de Kadt, D., & Wittels, S. B. (2019). Democratization and Economic Output in Sub-Saharan Africa. *Political Science Research and Methods*, 7(1), 63–84. <https://doi.org/10.1017/psrm.2016.15>
- Doudchenko, N., & Imbens, G. W. (2016). *Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis*. 22791. <https://doi.org/10.3386/w22791>
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, 7(1). <https://doi.org/10.1214/aos/1176344552>
- Eibl, F., & Hertog, S. (2024). From Rents to Welfare: Why Are Some Oil-Rich States Generous to Their People? *American Political Science Review*, 118(3), 1324–1343. <https://doi.org/10.1017/S0003055423000977>
- Funke, M., Schularick, M., & Trebesch, C. (2023). Populist Leaders and the Economy. *American Economic Review*, 113(12), 3249–3288. <https://doi.org/10.1257/aer.20202045>
- Gilens, M., Patterson, S., & Haines, P. (2021). Campaign Finance Regulations and Public Policy. *American Political Science Review*, 115(3), 1074–1081. <https://doi.org/10.1017/S0003055421000149>
- Hassell, H. J. G., & Holbein, J. B. (2024). Navigating Potential Pitfalls in Difference-in-Differences Designs: Reconciling Conflicting Findings on Mass Shootings' Effect on Electoral Outcomes. *American Political Science Review*, 1–21. <https://doi.org/10.1017/S0003055424000108>
- Hazlett, C., & Xu, Y. (2018). *Trajectory Balancing: A General Reweighting Approach to Causal Inference With Time-Series Cross-Sectional Data*. <https://doi.org/10.2139/ssrn.3214231>
- Hilbig, H., & Riaz, S. (2024). Natural Disasters and Green Party Support. *The Journal of Politics*, 86(1), 241–256. <https://doi.org/10.1086/726917>
- Hirshberg, D. A. (2022). *Synthdid: Synthetic difference in differences estimation*. <https://synth-inference.github.io/synthdid/>
- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945–960. <https://doi.org/10.1080/01621459.1986.10478354>
- Imai, K., & Kim, I. S. (2019). When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data? *American Journal of Political Science*, 63(2), 467–490. <https://doi.org/10.1111/ajps.12417>
- Imai, K., Kim, I. S., & Wang, E. H. (2023). Matching Methods for Causal Inference with Time-Series Cross-Sectional Data. *American Journal of Political Science*, 67(3), 587–605. <https://doi.org/10.1111/ajps.12685>

- Jackson, J. E. (2002). A Seemingly Unrelated Regression Model for Analyzing Multiparty Elections. *Political Analysis*, 10(1), 49–65. <https://doi.org/10.1093/pan/10.1.49>
- Katz, J. N., & King, G. (1999). A Statistical Model for Multiparty Electoral Data. *The American Political Science Review*, 93(1), 15–32. <https://doi.org/10.2307/2585758>
- Kellogg, M., Mogstad, M., Pouliot, G. A., & Torgovitsky, A. (2021). Combining Matching and Synthetic Control to Tradeoff Biases From Extrapolation and Interpolation. *Journal of the American Statistical Association*, 116(536), 1804–1816. <https://doi.org/10.1080/01621459.2021.1979562>
- Lal, A., Lockhart, M., Xu, Y., & Zu, Z. (2024). How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies. *Political Analysis*, 1–20. <https://doi.org/10.1017/pan.2024.2>
- Lipsmeyer, C. S., Philips, A. Q., Rutherford, A., & Whitten, G. D. (2019). Comparing Dynamic Pies: A Strategy for Modeling Compositional Variables in Time and Space. *Political Science Research and Methods*, 7(3), 523–540. <https://doi.org/10.1017/psrm.2017.39>
- Liu, L., Wang, Y., & Xu, Y. (2024). A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data. *American Journal of Political Science*, 68(1), 160–176. <https://doi.org/10.1111/ajps.12723>
- Meserve, S. A., & Pemstein, D. (2018). Google Politics: The Political Determinants of Internet Censorship in Democracies. *Political Science Research and Methods*, 6(2), 245–263. <https://doi.org/10.1017/psrm.2017.1>
- Miller, R. G. (1974). The Jackknife – A Review. *Biometrika*, 61(1), 1. <https://doi.org/10.2307/2334280>
- Mullahy, J. (2018). *Treatment Effects with Multiple Outcomes*. 25307. <https://doi.org/10.3386/w25307>
- Negi, A., & Wooldridge, J. M. (2024). Doubly Robust Estimation of Multivariate Fractional Outcome Means with Multivalued Treatments. *Econometric Reviews*, 43(2–4), 175–196. <https://doi.org/10.1080/07474938.2024.2310987>
- Pérez, V., Aybar, C., & Pavía, J. M. (2021). Spanish electoral archive. SEA database. *Scientific Data*, 8(1), 193. <https://doi.org/10.1038/s41597-021-00975-y>
- Philips, A. Q., Rutherford, A., & Whitten, G. D. (2015). The Dynamic Battle for Pieces of Pie—Modeling Party Support in Multi-Party Nations. *Electoral Studies*, 39, 264–274. <https://doi.org/10.1016/j.electstud.2015.03.019>
- Philips, A. Q., Rutherford, A., & Whitten, G. D. (2016). Dynamic Pie: A Strategy for Modeling Trade-Offs in Compositional Variables over Time. *American Journal of Political Science*, 60(1), 268–283. <https://doi.org/10.1111/ajps.12204>
- Pinotti, P. (2015). The Economic Costs of Organised Crime: Evidence from Southern Italy. *The Economic Journal*, 125(586), F203–F232. <https://doi.org/10.1111/eoj.12235>
- Robbins, M. W., Saunders, J., & Kilmer, B. (2017). A Framework for Synthetic Control Methods With High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-

- Specific Crime Intervention. *Journal of the American Statistical Association*, 112(517), 109–126. <https://doi.org/10.1080/01621459.2016.1213634>
- Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, 4(3), 305–322. <https://doi.org/10.1257/aeri.20210236>
- Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244. <https://doi.org/10.1016/j.jeconom.2023.03.008>
- Samartsidis, P., Seaman, S. R., Montagna, S., Charlett, A., Hickman, M., & De Angelis, D. (2021). A Bayesian Multivariate Factor Analysis Model for Evaluating an Intervention by Using Observational Time Series Data on Multiple Outcomes. *Journal of the Royal Statistical Society. Series A, (statistics in Society)*, 28(1), 155–166. <https://doi.org/10.1111/rssa.12569>
- Selb, P., & Munzert, S. (2018). Examining a Most Likely Case for Strong Campaign Effects: Hitler's Speeches and the Rise of the Nazi Party, 1927–1933. *American Political Science Review*, 112(4), 1050–1066. <https://doi.org/10.1017/S0003055418000424>
- Sun, L., Ben-Michael, E., & Feller, A. (2023). *Using Multiple Outcomes to Improve the Synthetic Control Method*. arXiv: 2311.16260. <https://doi.org/10.48550/arXiv.2311.16260>
- Thompson, D. M. (2023). Do Citizens Vote Against Incumbents Who Permit Local Immigration? Evidence from the Mariel Boatlift. *Political Science Research and Methods*, 11(4), 728–745. <https://doi.org/10.1017/psrm.2021.76>
- Timoneda, J. C., Escribà-Folch, A., & Chin, J. (2023). The Rush to Personalize: Power Concentration after Failed Coups in Dictatorships. *British Journal of Political Science*, 53(3), 878–901. <https://doi.org/10.1017/S0007123422000655>
- Tomz, M., Tucker, J. A., & Wittenberg, J. (2002). An Easy and Accurate Regression Model for Multiparty Electoral Data. *Political Analysis*, 10(1), 66–83. <https://doi.org/10.1093/pan/10.1.66>
- Trejo, G., & Nieto-Matiz, C. (2023). Containing Large-Scale Criminal Violence Through Internationalized Prosecution: How the Collaboration Between the CICIG and Guatemala's Law Enforcement Contributed to a Sustained Reduction in the Murder Rate. *Comparative Political Studies*, 56(9), 1328–1364. <https://doi.org/10.1177/00104140221139386>
- Xu, Y. (2017). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis*, 25(1), 57–76. <https://doi.org/10.1017/pan.2016.2>
- Xu, Y. (2024). Causal Inference with Time-Series Cross-Sectional Data: A Reflection. In J. M. Box-Steffensmeier, D. P. Christenson, & V. Sinclair-Chapman (Eds.), *Oxford Handbook of Engaged Methodological Pluralism in Political Science*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780192868282.013.30>

3.A Appendix

3.A.1 Formalization of Synthetic Control Methods

We consider a balanced panel with N units and T time periods. Where we denote units with $i \in \{1, \dots, N\}$ and periods with $t \in \{1, \dots, T\}$. The outcome of interest are K proportions, where each proportion value is denoted with $k \in \{1, \dots, K\}$. The outcome for unit i in period t is a multivariate vector of $Y_{it} = [y_{it1}, \dots, y_{itK}]$. With a proportion, it holds that $y_{itk} \in [0, 1]$ for all k and $\sum_{k=1}^K y_{itk} = 1$, forming a K -dimensional simplex.

We further have a binary treatment $D_{it} \in \{0, 1\}$. For (now) we assume that there is a share of control units that are never exposed to the treatment N_{co} , while there are units N_{tr} that are treated after time T_{pre} . For the analysis, we arrange the i units such that first all control units appear and then all treatment units $i \in \{1, \dots, N_{co}, N_{co} + 1, \dots, N\}$.

A DID method would estimate separate average treatment effects of exposure for each proportion, τ_k^{did} essentially solving a two-way fixed effects regression problem:

$$(\hat{\tau}_k^{did}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \right\} \quad (3.11)$$

for this specification, the (unweighted) average of the control units essentially serves as the counterfactual outcome for the treatment units after treatment exposure. The key identification parallel trends assumption makes this (unweighted) average a valid counterfactual outcome.

Synthetic methods instead calculate weights $\hat{\omega}_{ik}$ that align pre-exposure trends in the outcomes of the control with the treatment units, such that $\sum_{i=1}^{N_{co}} \hat{\omega}_{ik} y_{itk} \approx \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk}$ for $t \in \{1, \dots, T_{pre}\}$. The approach further specifies time-weights $\hat{\lambda}_{tk}$ that balance pre- and post-exposure periods for the different shares. The estimation then proceeds with finding a solution to the minimization problem.

$$(\hat{\tau}_k^{sdi}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=t}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \hat{\omega}_{ik} \hat{\lambda}_{tk} \right\} \quad (3.12)$$

As Arkhangelsky et al. point out, the synthetic difference-in-differences estimator can also be applied to classic synthetic method approaches (Arkhangelsky et al., 2021). Here the minimization problem excludes the time-weights and unit fixed effects.

Of central importance for SCMs is the calculation of the weights. In the application to proportions, these would be calculated for each proportion separately. Arkhangelsky et al. approach for this is an optimization problem with regularization:

$$(\hat{\omega}_{0k}, \hat{\omega}_k) = \arg \min_{\omega_{0k} \in \mathbb{R}, \omega_k \in \Omega} \ell_{uni}(\omega_{0k}, \omega_k) \text{ where} \quad (3.13)$$

$$\ell_{uni}(\omega_{0k}, \omega_k) = \sum_{t=1}^{T_{pre}} \left(\omega_{0k} + \sum_{i=1}^{N_{co}} \omega_{ik} y_{itk} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk} \right)^2 + \xi_k^2 T_{pre} \|\omega_k\|_2^2,$$

$$\Omega = \left\{ \omega_k \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_{ik} = 1, \omega_{ik} = \frac{1}{N_{tr}} \forall i \in \{N_{co}+1, \dots, N\} \right\}$$

where the regularization parameter is pre-specified as ξ

$$\xi_k = (N_{tr} T_{post})^{1/4} \hat{\sigma} \text{ with } \hat{\sigma}^2 = \frac{1}{N_{co}(T_{pre} - 1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} (\Delta_{itk} - \bar{\Delta}_k)^2, \quad (3.14)$$

$$\text{where } \Delta_{itk} = y_{it(t+1)k} - y_{ittk} \text{ and } \bar{\Delta}_k = \frac{1}{N_{co}(T_{pre} - 1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} \Delta_{itk} \quad (3.15)$$

The time weights are analogously implemented by solving:

$$(\hat{\lambda}_{0k}, \hat{\lambda}_k) = \arg \min_{\lambda_{0k} \in \mathbb{R}, \lambda_k \in \Lambda} \ell_{time}(\lambda_{0k}, \lambda_k) \text{ where} \quad (3.16)$$

$$\ell_{uni}(\lambda_{0k}, \lambda_{ik}) = \sum_{i=1}^{N_{co}} \left(\lambda_{0k} + \sum_{t=1}^{T_{pre}} \lambda_{tk} y_{itk} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T y_{itk} \right)^2,$$

$$\Lambda = \left\{ \lambda_k \in \mathbb{R}_+^T : \sum_{i=1}^{T_{pre}} \lambda_t = 1, \lambda_t = \frac{1}{T_{post}} \forall i \in \{T_{pre} + 1, \dots, T\} \right\}$$

The algorithm for synthetic Difference in Differences proceeds as follows:

- For each proportion $k \in \{1, \dots, K\}$:
 1. Compute regularization parameter ξ_k
 2. Compute unit weights $\hat{\omega}_k$
 3. Compute time weights $\hat{\lambda}_k$
 4. Compute the SDID estimator via the weighted DID regression

$$(\hat{\tau}_k^{sdid}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \hat{\omega}_{ik} \hat{\lambda}_{tk} \right\} \quad (3.17)$$

3.A.2 Proofs

Lemma 1: By the linearity of expectation, we can sum the treatment effects across all categories k : $\sum_{k=1}^K \tau_k = E \left[\sum_{k=1}^K y_{ik}(1) - \sum_{k=1}^K y_{ik}(0) \mid D_i = 1 \right]$. Next, we apply the compositional constraint on proportional outcomes, which states that for both treated and control units: $\sum_{k=1}^K y_{ik}(1) = 1$ and $\sum_{k=1}^K y_{ik}(0) = 1$. Thus, the sum of the potential outcomes under both treatment and control conditions is always 1 for each observation. Substituting this into the equation, we get: $\sum_{k=1}^K \tau_k = E [1 - 1 \mid D_i = 1] = E [0 \mid D_i = 1]$. Finally, since the expectation of zero is zero, we conclude that: $\sum_{k=1}^K \tau_k = 0$.

Theorem 1: We begin with the compositional constraints: $\sum_{k=1}^K y_{itk} = 1$ and $\sum_{k=1}^K \bar{y}_{i.k} = \frac{1}{T_0} \sum_{t=1}^{T_0} \sum_{k=1}^K y_{itk} = 1$. Additionally, we have $\sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk} = 0$, which can be shown by expanding the terms: $\sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk} = \sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} (y_{jtk} - \bar{y}_{i.k}) = \sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} y_{jtk} - \sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} \bar{y}_{i.k}$. Since the weights w_{jk} are common across outcome categories k , we can simplify: $\sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} y_{jtk} = \sum_{j=1}^{N_{co}} w_j$ and $\sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} \bar{y}_{i.k} = \sum_{j=1}^{N_{co}} w_j \sum_{k=1}^K \bar{y}_{i.k} = \sum_{j=1}^{N_{co}} w_j$. This leaves us with: $\sum_{k=1}^K \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk} = \sum_{j=1}^{N_{co}} w_j - \sum_{j=1}^{N_{co}} w_j = 0$. Now, we apply these results to the sum of the ATT estimates:

$$\sum_{k=1}^K \hat{\tau}_k = \frac{1}{N_{tr}} \sum_{k=1}^K \left(\sum_{i=N_{co}+1}^N y_{iT_k} \right) - \frac{1}{N_{tr}} \sum_{k=1}^K \left(\sum_{i=N_{co}+1}^N \left(\bar{y}_{i.k} + \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk} \right) \right).$$

The first term for the treated group: $\frac{1}{N_{tr}} \sum_{k=1}^K \left(\sum_{i=N_{co}+1}^N y_{iT_k} \right) = 1$, since $\sum_{k=1}^K y_{iT_k} = 1$ for each treated unit. We expand the second term for the control group as: $\frac{1}{N_{tr}} \sum_{k=1}^K \sum_{i=N_{co}+1}^N \bar{y}_{i.k} + \frac{1}{N_{tr}} \sum_{k=1}^K \sum_{i=N_{co}+1}^N \sum_{j=1}^{N_{co}} w_{jk} \dot{y}_{jtk}$. The first part simplifies to 1, since $\sum_{k=1}^K \bar{y}_{i.k} = 1$. The second part is 0, as shown earlier. Thus, we have: $\sum_{k=1}^K \hat{\tau}_k = 1 - 1 - 0 = 0$.

Example without and with constant control comparisons: Consider $K = 2$ with one treated unit and two control units. The treated unit has $y_{iT_1}(1) = 0.6$, $y_{iT_2}(1) = 0.4$. The control units have: $y_{1T_1}(0) = 0.7$, $y_{1T_2}(0) = 0.3$, $y_{2T_1}(0) = 0.5$, $y_{2T_2}(0) = 0.5$.

We first show the case with unequal weights. The weights are $w_{11} = 0.8$, $w_{21} = 0.2$ for $k = 1$, and $w_{12} = 0.2$, $w_{22} = 0.8$ for $k = 2$. The pre-treatment averages are $\bar{y}_{i.1} = 0.7$ and $\bar{y}_{i.2} = 0.3$. The demeaned outcomes for controls are: $\dot{y}_{11} = 0$, $\dot{y}_{21} = -0.2$, $\dot{y}_{12} = 0$, $\dot{y}_{22} = 0.2$. The ATT for $k = 1$: $\hat{\tau}_1 = 0.6 - (0.7 - 0.04) = 0.6 - 0.66 = -0.06$, and for $k = 2$: $\hat{\tau}_2 = 0.4 - (0.3 + 0.16) = 0.4 - 0.46 = -0.06$. The sum of the treatment effects is: $\sum_{k=1}^2 \hat{\tau}_k = -0.12$, which violates the constraint $\sum_{k=1}^K \hat{\tau}_k = 0$ due to unequal weights.

Now assume equal weights for the outcome categories but different weights across units: $w_{11} = w_{12} = 0.5$ for Control 1, $w_{21} = w_{22} = 0.5$ for Control 2. The pre-treatment averages are $\bar{y}_{i.1} = 0.7$ and $\bar{y}_{i.2} = 0.3$. The demeaned outcomes for controls are: $\dot{y}_{11} = 0$, $\dot{y}_{21} = -0.2$, $\dot{y}_{12} = 0$, $\dot{y}_{22} = 0.2$. The ATT for $k = 1$ is: $\hat{\tau}_1 = 0.6 - (0.7 + 0.5 \cdot 0 + 0.5 \cdot (-0.2)) = 0.6 - 0.6 = 0$, and for $k = 2$: $\hat{\tau}_2 = 0.4 - (0.3 + 0.5 \cdot 0 + 0.5 \cdot 0.2) = 0.4 - 0.4 = 0$. Thus, the sum of the treatment effects is: $\sum_{k=1}^2 \hat{\tau}_k = 0 + 0 = 0$. With equal weights for categories, the compositional constraint $\sum_{k=1}^K \tau_k = 0$ is satisfied.

3.A.3 Algorithm for Synthetic control methods with common weights

Algorithm 2. Compute the Synthetic Control (SC) Estimator

- 1: Compute regularization parameter ξ
- 2: Compute unit weights $\hat{\omega}$ solving minimization:

$$(\hat{\omega}) = \arg \min_{\hat{\omega} \in \mathbb{R}, \hat{\omega} \in \Omega} \ell_{uni}(\omega) \text{ where} \quad (3.18)$$

$$\ell_{uni}(\omega) = \sum_{k=1}^K \left(\sum_{t=1}^{T_{pre}} \left(\sum_{i=1}^{N_{co}} \omega_i y_{itk} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N y_{itk} \right)^2 \right) + \xi^2 T_{pre} \|\omega\|_2^2,$$

$$\Omega = \left\{ \omega \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_i = 1, \omega_i = \frac{1}{N_{tr}} \forall i \in \{N_{co} + 1, \dots, N\} \right\}$$

- 3: **for** each proportion $k \in \{1, \dots, K\}$ **do**
- 4: Compute the SC estimator via the weighted regression:

$$(\hat{\tau}_k^{sdid}, \hat{\mu}_k, \hat{\alpha}_{ik}, \hat{\beta}_{tk}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (y_{itk} - \mu_k - \alpha_{ik} - \beta_{tk} - D_{it}\tau)^2 \hat{\omega}_i \right\} \quad (3.19)$$

- 5: **end for**
-

3.A.4 Simulation

3.A.4.1 Data generating process

We use a dynamic latent state model to set up the data generating process for the simulation. We generate balanced panel data with T time points, N units, and K proportions. As mentioned above, $i \in \{1, \dots, N\}$ refers to units, $t \in \{0, \dots, T\}$ refers to time points, and $k \in \{1, \dots, K\}$ refers to proportions. N_1 treatment units are treated from time point T_1 onward.¹⁶

First, we initialize latent unit factors (state variables) for each unit i and proportion k before the panel starts at $t = 0$ by taking draws from independent normal distributions:

¹⁶There is no sequential exposure to treatment and no treatment reversal, although these processes can easily be integrated.

$\gamma_{i,k,0} \sim \mathcal{N}(\mu_{\gamma_{k,0}}, \sigma_{\gamma_{k,0}})$, $v_{i,k,0} \sim \mathcal{N}(\mu_{v_{k,0}}, \sigma_{v_{k,0}})$. γ describes the local levels for the different proportions and v describes the local trends. The hyperparameters $\mu_{\gamma_{k,0}}$, $\sigma_{\gamma_{k,0}}$, $\mu_{v_{k,0}}$, and $\sigma_{v_{k,0}}$ allow us to set initial differences between units. As we see later, the local trends model is useful for setting time processes that violate parallel trends assumptions.

Units are allocated to treatment and control probabilistically based on a linear combination of the latent factors of one of the proportions j and a parameter δ that governs whether the treatment is selected based on local level differences or local trend differences. We define: $L_i = \psi\gamma_{i,j,0} + (1 - \psi)v_{i,j,0}$. Based on L_i , we calculate treatment assignment probabilities using a logistic function with parameters λ_0 and λ : $\pi_i = \frac{1}{1 + \exp(-\lambda_0 - \lambda_1 L_i)}$. We draw treatment assignment from a Bernoulli distribution: $D_i \sim \text{Bernoulli}(\pi_i)$. Here, λ_1 is a hyperparameter that sets the strength of treatment selection based on latent factors.¹⁷ Based on this, we set the treatment indicator $W_{i,t}$ for unit i at time t to 1 if $D_i = 1$ and $t \geq T_1$, otherwise, zero.

To simulate the observed outcomes for different time points, we update the latent unit factors and time factors dynamically over time t :

$$\begin{aligned}\gamma_{i,k,t+1} &= \gamma_{i,k,t} + v_{i,k,t} + \epsilon_{\gamma_{i,k,t}}, \quad \text{with } \epsilon_{\gamma_{i,k,t}} \sim \mathcal{N}(0, \sigma_{\gamma_k}) \\ v_{i,k,t+1} &= v_{i,k,t} + \epsilon_{v_{i,k,t}}, \quad \text{with } \epsilon_{v_{i,k,t}} \sim \mathcal{N}(0, \sigma_{v_k})\end{aligned}$$

For each time t and factor k , the outcome $Y_{i,t}^k$ for unit i is generated as:

$$\tilde{Y}_{i,k,t} = \gamma_{i,k,t} + \tau_k W_{i,t} + \delta_{k,t} + \epsilon_{i,k,t}$$

¹⁷To ensure that the number of treated units in our dataset is exactly N_1 , we first set $\lambda_0 = \log\left(\frac{N_1}{N}\right)$. Then, we adjust the assignment such that if the number of treatment assignments exceeds the target, we resample:

- If $\sum D_i > N_1$, select N_1 units from the assigned treated units.
- If $\sum D_i < N_1$, select $(N - N_1)$ units from the assigned control units.
- If $\sum D_i = 0$, select N_1 units randomly from all units.

The final assignment ensures exactly N_1 treated units.

where $\epsilon_{i,t} \sim \mathcal{N}(0, \sigma_\epsilon)$ are idiosyncratic errors and $\delta_{k,t} \sim \mathcal{N}(0, \sigma_\delta)$ are common time effects. Based on these, we simulate the proportional outcomes using a multinomial link:

$$Y_{i,t,k} = \frac{\exp(\tilde{Y}_{i,t,k})}{\sum_{k=1}^K \exp(\tilde{Y}_{i,t,k})}$$

The multinomial link ensures that the sum of outcome proportions for each time point is defined on the simplex, i.e., $\sum_{k=1}^K Y_{i,t,k} = 1$.

With this data generating process, we can draw samples of compositional panel data. We can vary the time-horizon, the treatment process, the number of units and the local trend and level dynamics, and the type of treatment selection that induces violation of the parallel trends' assumption.

3.A.4.2 Set-up

In the main text simulation, we vary parameters to assess method robustness under diverse scenarios. We examine time-series lengths of 5 and 10 periods, with treatment applied in either the last period or the last two. Sample sizes range across 50, 200, and 1,000 units, with treatment proportions of 10% and 50%. Treatment assignment is based on either local levels, simulating initial differences, or local trends, introducing diverging pre-treatment trends.

We further fix the other parameters of the DGP described in SM 3.A.4.1. All prior local levels parameters are set to $\sigma_{\gamma_{k,0}} = 1$ and $\mu_{\gamma_{k,0}} = 0$. For prior local trends of the first proportion, we impose $\sigma_{v_{1,0}} = 0.5$ and $\mu_{v_{1,0}} = 0$. All other prior local trends are set to zero. The variability of the dynamics is set to $\sigma_{\gamma_k} = 0.1$ and $\sigma_{v_k} = 0.2$. The standard deviation of the errors is set to $\sigma_\epsilon = 0.1$ and $\sigma_\delta = 0.2$. The treatment selection parameter λ_1 is set to 2.

3.A.4.3 Additional results

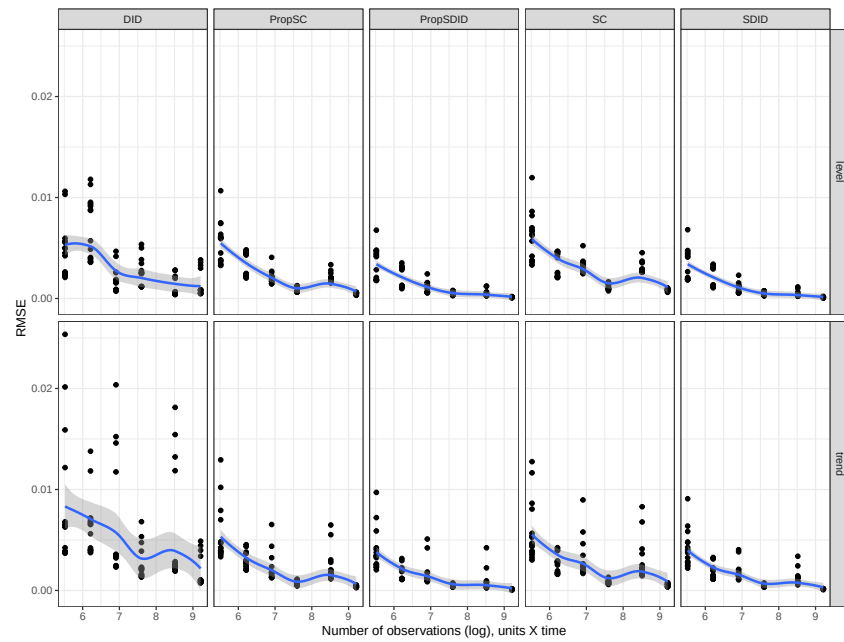


Figure 3.3. Monte Carlo evaluation of difference-in-differences (DID), synthetic difference-in-differences (SDID), proportional synthetic difference-in-differences (PropSDID), synthetic control (SC), and proportional synthetic control (PropSC). The figure presents the Root Mean Square Error (RMSE) across different simulation configurations, with case numbers represented as the natural logarithm of the product of time points and units.

