How Mainstream Politicians Erode Norms:

Evidence from two survey experiments

Vicente Valentim[1], Daniel Ziblatt[2], Elias Dinas[3]

Abstract: What is the effect of xenophobic rhetoric by mainstream politicians on norms of tolerance? How does this compare to similar statements made by radical-right politicians? In two survey experiments, we find that statements made by mainstream politicians are more likely to lead to norm erosion than those made by radical-right politicians. Subsample analyses suggests that this is because statements by radical-right politicians generate backlash among left-wing individuals, who update their norm perceptions upward. This backlash effect is no longer noticeable when similar statements are made by mainstream right politicians. This difference may be due to mainstream politicians representing the views of a larger part of the population, or having higher status. Our results highlight the pivotal role of mainstream politicians in enforcing or eroding democratic norms. They also show that similar political statements have different effects depending on their sender.

[1] Nuffield College, University of Oxford.
[2] Harvard University and WZB Berlin Social Science Center.
[3] European University Institute.

Introduction

There is a gap between theoretical and empirical work on democratic backsliding. The former places emphasis on the role of social norms in sustaining democratic stability (Levitsky and Ziblatt 2018) while the latter has only recently begun to explore the ways in which such norms can change. One key question in the empirical literature is how the success of non-mainstream political actors can alter people's perceptions as to how acceptable their opinions are. For example, Bursztyn et al. (2020) show that when people perceive others around them as supporting Trump, they become more likely to express their own openly racist attitudes. Similarly, supporters of radical-right parties are more likely to acknowledge their political preferences in surveys after these parties marginally cross thresholds of parliamentary representation (Valentim 2021). This recent empirical work's focus on radical political actors contrasts with a long-standing theoretical scholarship which contends that whether mainstream political actors abide by established norms is crucial for democratic resilience (Linz, 1978; Levitsky and Ziblatt 2018).

We try to bridge this gap between empirical and theoretical work. To that end, we focus on the way *mainstream* politicians can also affect norm change when they use xenophobic rhetoric.[4] There are two pathways by which mainstream politicians may feel impelled to use xenophobic rhetoric with potential norm breaching effects. First, mainstream politicians may themselves introduce xenophobic rhetoric, sometimes pre-emptively to block the rise of the radical right. Second, mainstream politicians may deploy the issues that the far right has already put in the agenda (Meguid 2005; Green-Pedersen and Mortensen 2010; Meyer and Rosenberger 2015; Gessler and Hunger 2022), and may even adopt their policy stances (Rooduijn et al. 2014; Abou-Chadi and Krause 2017). It is surprising that although these two dynamics have been widely acknowledged, no work thus far has investigated their implications for norm perceptions. What happens when mainstream politicians themselves make xenophobic statements? How does the effect of those statements compare to similar statements made by radical-right politicians?

Answering these questions empirically is difficult because it requires manipulating the source of a message without altering its content. One possibility would be to present experimental subjects with identical fictional statements, varying their partisan origins. This strategy, however, suffers from ethical concerns generated from deception. Moreover, it raises issues of external and internal validity due to credibility issues: some parties are more natural sources for xenophobic statements than others.

[4] We follow Pop-Eleches's (2010, p. 225) definition of mainstream parties as those that "represent an ideological orientation that can be mapped with reasonable accuracy onto the mainstream ideological spectrum of established Western democracies".

To tackle these concerns, we run two pre-registered survey experiments that leverage near-identical xenophobic statements made by radical-right and center-right politicians in Germany. This allows us to compare reactions to similar statements made by the two types of politicians without engaging in deception. We use a battery of outcomes that allow us to tap into several components of anti-prejudice norms: perceptions about how widespread, as well as appropriate, xenophobic beliefs are; willingness to violate the injunctive norm by publicly expressing xenophobic sentiments; and willingness to sanction openly expressed xenophobic preferences.

Our results suggest that exposure to xenophobic cues from mainstream right politicians leads to a more across-the-board perception of norm erosion than when similar statements are made by radical-right politicians. Subsample analyses suggest that this is because radical-right statements generate backlash among left-wing voters, who report stronger norm perceptions against prejudice. This backlash effect, however, is not found when the norm-breaching statement is made by politicians from mainstream parties. In that case, even left-wing voters perceive the norm to have eroded.

We speak to three strands in the existing literature. First, we enrich a nascent literature on how political actors affect outcomes not only via persuasion, priming, and framing, but rather by changing people's perceptions about how widely acceptable given attitudes or opinions are. While previous work has focused on the impact of leaders and parties on shaping people's attitudes, their impact on social norms remains still largely overlooked.

Second, we bridge a literature that focuses on party competition with a more normative literature on democratic backsliding. For the most part, the success of the radical right has been seen as a threat to democracy either because of the fear that these policies get implemented or that they gain more public support. We provide evidence for the next step in this process: the multiplier effect of such change in political stances among the mainstream right on norm erosion.

Finally, we provide empirical backing to the theoretical claim that the extent to which mainstream politicians give in to norm-breaching behavior and rhetoric is crucial for the sustainment or erosion of democratic norms.

Research Design

We test the relative effect of norm-breaching statements by radical-right and mainstream politicians using two pre-registered experiments[5] — a coordination game and a third-party punishment experiment — embedded within an

[5] An anonymous version of the pre-registration can be found in Appendix G.

original and nationally representative survey fielded in Germany. Fieldwork took place between January 20 and February 1, 2022. All participants provided informed consent before they participated in the study.

We chose Germany for two reasons. First, it is a context where there is still a strong and widespread anti-prejudice norm (Blinder et al. 2013), thus allowing us to study the effect of treatment with xenophobic elite cues on norm perception. Second, studying Germany allows us to take advantage of the increasing popularity of the radical right party (the AfD), as well as the news of the US withdrawal from Afghanistan and worries about a second migration crisis. These events led some center-right politicians (from the CDU) to adopt anti-immigrant rhetoric ahead of the 2021 German federal elections. This setting, together with the high salience of the issue of migration, provided a realistic context for the anti-immigration quotes made by CDU politicians that we presented to respondents.

To build our treatments, we identified quotes from CDU and AfD politicians that are roughly identical in content and intensity. These quotes were circulated to a number of independent coders to see if they were able to distinguish them from each other or if the quotes could be attributed to the original speaker. We only used quotes that passed this test. Figure B.2 in the Online Appendix provides evidence of how coders coded the quotes.

Using complete random assignment, participants were divided into five distinct treatment conditions — Radical Right Party (*RRP*) *approve*; *Mainstream approve*; *Mainstream approve* + *RRP approve*; *Mainstream disapprove* + *RRP approve*; and *Control*. Participants in each of these conditions received the same survey with two embedded experiments, though they were exposed to different treatment vignettes according to the condition that they were randomized into.

In the first two conditions (*RRP approve, Mainstream approve)*, participants were exposed to vignettes which consist of anti-immigrant quotes. These quotes attributed to an unnamed German radical right or mainstream right politician, respectively. In the third condition (*Mainstream approve* + *RRP approve)*, participants were exposed to a similar vignette where anti-immigrant quotes are attributed to both an unnamed German radical right and mainstream right politician. The fourth vignette (*Mainstream disapprove* + *RRP approve*) contains a pro-immigration quote from a German mainstream right politician, along with anti-immigration quotes from a radical right politician. The fifth treatment condition (*Control)* is a similar vignette on an unrelated issue, specifically, food waste reductions. Table 1 summarizes the treatment arms, while Table C.1 in the Online Appendix shows each vignette in detail.

Table 1: Treatment Conditions, approving refers to an anti-immigrant statement, while disapproving refers to a pro-immigrant statement.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	Mainstream Approves	RRP Approves	Both Approve	Mainstream disapproves RRP Approves	Control
Mainstream Right	+		+	-	
RRP Right		+	+	+	

Note: (+) indicates approval; (-) indicates disapproval.

As outcomes, we examine three measures of norm perception. These measures build on Bicchieri's (2016) definition of social norm as the product of both empirical expectations (others share this view) and normative expectations (others believe this view is appropriate).

1) Perception about the share of other participants willing to express xenophobic sentiments (empirical expectations), and share of others who would find such expression appropriate (normative expectations). Both variables were measured on a scale 1-100 and incentivized so that respondents who got closer to the real number would be eligible for a bonus payment.

2) Norm conformity, measured in terms of willingness to violate the injunctive norm by publicly expressing xenophobic sentiments. This variable and the previous one are measured using the coordination game, which relies on a hypothetical petition to limit immigration from Afghanistan and the Middle East that respondents have the option of publicly signing. This is shown in Figure A.1 in the Online Appendix. To avoid concerns with deception, respondents were explicitly told that the petition was hypothetical.

3) Willingness to sanction others who publicly express xenophobic sentiments. Participants were presented with a Twitter post containing anti-immigrant slurs from an anonymous user (shown in Figure A.2 in the Online Appendix). They were informed that these comments had been flagged as potential violations of Twitter's hate speech policy, and provided the option of reporting similar comments. To avoid deception, the Twitter post was adapted from an actual message that was removed by Twitter administrators for violating Twitter's hate speech policy.

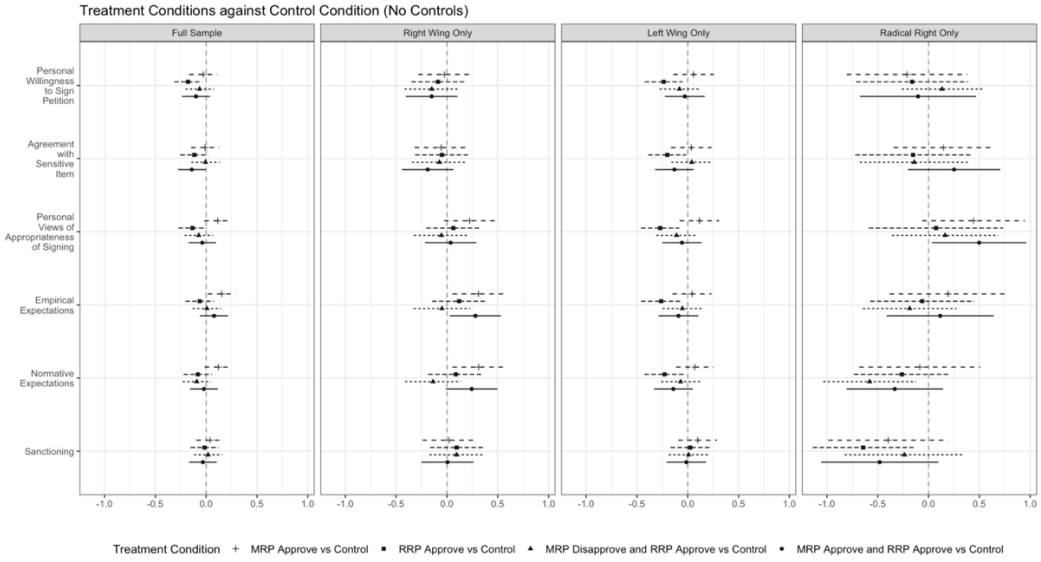
After exclusions, the final sample contained 2009 respondents. Respondents were excluded if they failed to pass a pre-treatment attention check, or if they did not complete the survey in its entirety. We have examined the

breakdown of responses that were excluded and found no evidence of differential attrition. This evidence is provided in Appendix D. The final sample is nationally representative by age and gender.[6]

Results

Figure 1 shows the results of exposure to each of the treatment arms. As per pre-registration, the coefficients report results from linear regression models that compare the values of each of our six outcome variables in each treatment arm to those found in the control group (which was not exposed to any norm-breaching statement). The Figure reports all outcomes included in the survey and specified in the pre-analysis plan.[7]

Figure 1: Effect of each treatment condition on outcome variables (95% CIs, standardized coefficients).



The results using the whole sample (first facet column) suggest that exposure to xenophobic cues from radical-right politicians (represented by square-shaped coefficients) does not erode perception of the norm against prejudice on average. Instead, it appears to generate a backlash effect that strengthens norm perception — when individuals are asked about their willingness to sign the petition, individuals in the *RRP approve* condition are less willing to sign the petition, less likely to agree with the sensitive item, and view signing the petition as less normatively acceptable.

The figure is very different when it is mainstream right—not radical-right—politicians who make the norm-breaching statement. Indeed, as shown by the cross-shaped coefficients, when mainstream right politicians make

[6] The study received approval from the [Identifying Information] research ethics committee (2021/0/117). An anonymized version of the consent form for participation in the study can be found in Appendix H.
[7] Tables E1-E6 of Online Appendix E present the full results from the analyses giving rise to the estimates shown in Figure.

such statements, individuals view signing the petition as more appropriate, they think that a higher number of others will sign the petition (empirical expectations) and also think that a higher number of others will deem it acceptable to sign (normative expectations).

In our pre-analysis plan, we were interested in understanding the effects of norm erosion statements by both radical-right and mainstream politicians. We expected, however, that both would lead to norm erosion.

To understand what drives this difference in the magnitude and direction of the effect, we look into the subsample analyses split on individual's party attachment. Note that these are exploratory analyses that we had specified in the pre-analyses plan, but that we had no particular expectation regarding them. The reason we look into them is to understand what is driving the difference in the effects depending on who makes the norm-eroding claim (radical-right or mainstream politicians).

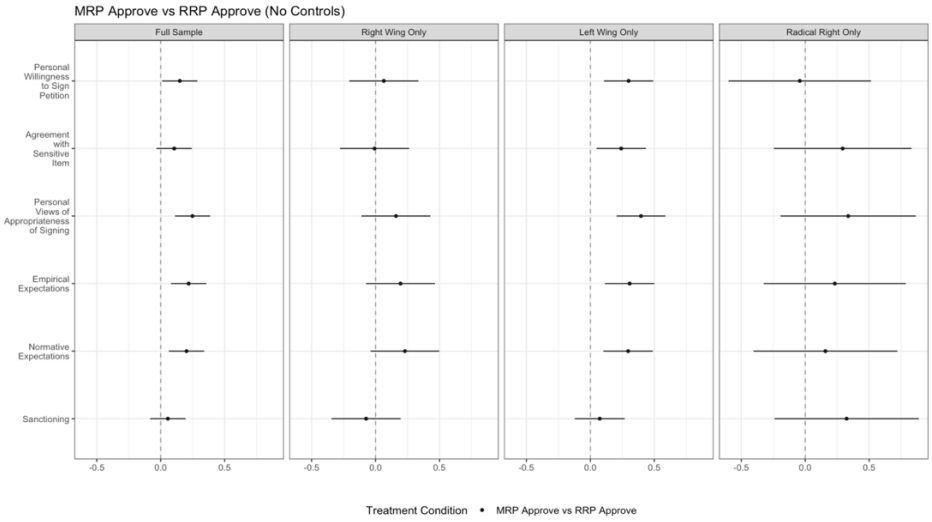
These analyses are shown in the second to fourth facet column. Each represents a different subsample: full sample, right-wing respondents (those who support FPD or CDU/CSU), left-wing respondents (those who support SPD, Greens, or Die Linke), and radical-right respondents (those who support AfD). These subsample analyses provide important insights into the reason why we find different effects across these two treatments. Left-wing individuals (third column) who are exposed to a norm-breaching statement by a radical-right politician seem to become increasingly supportive of the established norm; and to perceive it as stronger (square-shaped coefficients). However, this backlash effect is not found when the statement is made by a mainstream right, instead of radical-right politician. If anything, in response to this treatment (MRP approve), left-wing individuals react in the opposite direction—even if many coefficients fail to reach statistical significance. When it comes to right-wing individuals (second column), we also find some evidence of a movement in the same direction. These effects, however, are smaller and, for the most part, far from statistical significance.

The Online Appendix provides some robustness checks and additional analyses. Figures F.1 and F.2 replicate Figure 1 and Figure 2, respectively, after adding controls. The results remain substantively similar. Figures F.3 to F.11 show that there are no unexpected heterogeneous effects when the samples are split in terms of gender, income and education. Likewise, the findings are generally not sensitive to the removal of each control (Figures F.12 to F.15) or the removal of respondents from any particular state, with the exception of Nordrhein-Westfalen (Figures F.16 to F.31).

This finding is in line with the expectation put forward by previous literature (e.g., Bischof and Wagner, 2019), according to which left-wing individuals should react to radical-right norm erosion by holding closer to the very norms that are breached by those actors. However, the findings also suggest that such is only the case when the norm-eroding actor is a radical-right politician. When mainstream right politicians make similar statements, norm erosion becomes much more likely because it does not face backlash by left-wing individuals.

To provide a more direct comparison of the effect of norm breaching statements by radical-right and mainstream right politicians, we compare those treatment arms against one another—instead of comparing both against the control condition, as in Figure 1.

Figure 2: Comparison of MRP approve against RRP approve (95% CIs, standardized coefficients).



The results are shown in Figure 2.[8] The Figure makes clear how the party membership of the politician making the norm breaching claim drastically changes the probability of norm erosion. Exposure to the *Mainstream Approve* condition makes the full sample much more likely to report willingness to sign the petition, agree with the sensitive item, view it as appropriate, expect that others will also sign the petition, and expect that others will find it appropriate to sign it. In line with the previous findings, this differential response to the two treatments is particularly pronounced for left-wing individuals. In line with the findings of Figure 1, we also find some differences in right-wing individuals, but these are smaller and, for the most part, far from statistical significance.

[8] Tables E7 to E11 of Online Appendix provides the full results from the analyses giving rise to the estimates reported in Figure 2.

Discussion: The pivotal role of mainstream politicians in norm erosion

The previous section showed that norm-breaching statements by mainstream politicians erode democratic norms more than do statements by radical-right politicians. Based on previous literature, we argue that there are two likely reasons why norm change is more widespread when mainstream politicians themselves engage in norm-breaching rhetoric. First, these politicians have wider electoral support. For norms to change in society at large, individuals need to perceive that a tipping point has been reached (Andreoni et al., 2021) and that a sufficiently large number of individuals in their reference network opposes the norm in place (Bicchieri, 2016). Since mainstream parties are typically larger than non-mainstream parties, when they make norm breaching statements, voters are more likely to update their perceptions of how many others support the established norm.

A second reason may simply be that norm conformity increases with the social status of the party. Social status plays a crucial role insofar as individuals abide by established norms and are punished for deviating from them. Previous research has shown that conformity increases with group status (Asch 1956), that individuals that occupy high status positions are less likely to be sanctioned for norm deviations (Hoff, Kshetramade and Fehr, 2011) and that the motivation to impress those with higher status is a crucial mechanism driving conformity (Kim, Kim, and Kim 2021).

Our results may be explained by a similar mechanism at the party level. Mainstream parties occupy a more prestigious position as guardians of democratic norms because they often cite themselves as sharing values and norms of the establishment (Valentim and Widmann 2021), enjoy more long-term ties with the electorate (Converse and Pierce 1992), and are thus considered institutional pillars of the system (Converse 1969). As a consequence, they are also better suited to change these norms than parties entering the system as *pariahs* (Van Spanje and Van Der Brug 2007).

Empirically testing which of these mechanisms is driving our effects is beyond the scope of our study, and a question that subsequent research may try to answer. Still, we believe these explanations provide theoretical reasons that help understand why mainstream politicians making norm eroding statements may represent a crucial step for norm change to reach the tipping points that lead to a more generalized norm change than similar statements made by radical-right politicians.

Conclusion

The results of the two survey experiments we report highlight how mainstream right politicians play a crucial role in the enforcement of democracy and its associated norms (Ziblatt 2017). Their statements have a stronger effect on mass-level democratic norms such as norms against xenophobia than do statements by radical-right politicians. This finding has important implications. Mainstream politicians may be tempted to give in to radical-right discourse as a ways of reacting or preventing their success. Our results suggest, however, that such accommodative strategies (Meguid 2008) may have detrimental effects for democracy. This is even more the case since, taking the findings from previous research into consideration, there seems to be no trade-off between keeping democratic norms in place and electoral success. As previous studies have shown, accommodative strategies also do not pay off electorally for the mainstream right (Spoon and Kluver 2020; Krause et al forthcoming).

Our results also suggest that the sender of a political message matters more than the actual content of that message. The results of our experiment suggest that a similar statement can have very different effects depending on the actor that makes it, and how that actor is perceived by voters. In so doing, our results are in line with results in recent research on social media which has shown that, in that realm, the identity of the sender significantly affects the way other users react to it (Taylor et al., 2022).

At the same time, this finding opens avenues for future research. Future studies may explore whether this finding travels to other issues or dyads of politicians. For example, is the finding specific to norm-breaching statements, or does it travel to statements that are not at odds with social norms? Does it matter that the statement is on migration, a topic over which radical-right politicians have ownership? Does it matter that, in our setting, both politicians are placed in the same ideological camp (right)? Answering these questions would allow for a better understanding of our main finding that the sender of a political message matters more than the content.

References

Abou-Chadi, Tarik, and Werner Krause. 'The Causal Effect of Radical Right Success on Mainstream Parties' Policy Positions: A Regression Discontinuity Approach'. *British Journal of Political Science*, 2018, 1–19.

Andreoni, James, Nikos Nikiforakis, and Simon Siegenthaler. 'Predicting Social Tipping and Norm Change in Controlled Experiments'. *Proceedings of the National Academy of Sciences* 118, no. 16 (2021).

Asch, Solomon E. 'Studies of Independence and Conformity: I. A Minority of One against a Unanimous Majority.' *Psychological Monographs: General and Applied* 70, no. 9 (1956): 1.

Bicchieri, Cristina. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York, NY: Oxford University Press., 2016.

Bischof, Daniel, and Markus Wagner. "Do Voters Polarize When Radical Parties Enter Parliament?" *American Journal of Political Science* 63, no. 4 (2019): 888-904.

Blinder, Scott, Robert Ford, and Elisabeth Ivarsflaten. 'The Better Angels of Our Nature: How the Antiprejudice Norm Affects Policy and Party Preferences in Great Britain and Germany'. *American Journal of Political Science* 57, no. 4 (2013): 841–57.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin. "From Extreme to Mainstream: The Erosion of Social Norms." *American Economic Review* 110, no. 11 (2020): 3522-548.

Clayton, Katherine, Nicholas T. Davis, Brendan Nyhan, Ethan Porter, Timothy J. Ryan, and Thomas J. Wood. "Elite Rhetoric Can Undermine Democratic Norms." *Proceedings of the National Academy of Sciences* 118, no. 23 (2021).

Converse, Philip E. 'Of Time and Partisan Stability'. *Comparative Political Studies* 2, no. 2 (1969): 139–71.

Converse, Philip E., and Roy Pierce. 'Partisanship and the Party System'. *Political Behavior* 14 (1992): 239–59.

Gessler, Theresa, and Sophia Hunger. 'How the Refugee Crisis and Radical Right Parties Shape Party Competition on Immigration'. *Political Science Research and Methods* 10, no. 3 (2022): 524–44.

Green-Pedersen, Christoffer, and Peter B. Mortensen. 'Who Sets the Agenda and Who Responds to It in the Danish Parliament? A New Model of Issue Competition and Agenda-Setting'. *European Journal of Political Research* 49, no. 2 (2010): 257–81.

Hjorth, Frederik, and Martin Vinæs Larsen. 'When Does Accommodation Work? Electoral Effects of Mainstream Left Position Taking on Immigration'. *British Journal of Political Science* 52, no. 2 (2022): 949–57.

Hoff, Karla, Mayuresh Kshetramade, and Ernst Fehr. 'Caste and Punishment: The Legacy of Caste Culture in Norm Enforcement'. *The Economic Journal* 121, no. 556 (2011): F449–75.

Kim, Daeeun, JuYoung Kim, and Hackjin Kim. 'Increased Conformity to Social Hierarchy Under Public Eyes'. *Frontiers in Psychology* 12 (2021): 636801.

Kingzette, Jon, James N. Druckman, Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. "How Affective Polarization Undermines Support for Democratic Norms." *Public Opinion Quarterly*, 2021.

Krause, Werner, Denis Cohen, and Tarik Abou-Chadi. 'Does Accommodation Work? Mainstream Party Strategies and the Success of Radical Right Parties'. *Political Science Research and Methods*, forthcoming.

Levitsky, Steven, and Daniel Ziblatt. *How Democracies Die*. London: Penguin, 2019.

Linz, Juan J. 'Non-Competitive Elections in Europe'. In Elections Without Choice, edited by Guy Hermet, Richard Rose, and Alain Rouquié, 36–65. London: Palgrave Macmillan UK, 1978.

Meguid, Bonnie M. *Party Competition between Unequals: Strategies and Electoral Fortunes in Western Europe*. Cambridge: Cambridge University Press, 2008.

Meléndez, Carlos, and Cristóbal Rovira Kaltwasser. 'Negative Partisanship towards the Populist Radical Right and Democratic Resilience in Western Europe'. *Democratization* 28, no. 5 (2021): 949–69.

Meyer, Sarah, and Sieglinde Rosenberger. 'Just a Shadow? The Role of Radical Right Parties in the Politicization of Immigration, 1995–2009'. *Politics and Governance* 3, no. 2 (2015): 1–17.

Mudde, Cas. 'The Populist Zeitgeist'. *Government and Opposition* 39, no. 4 (2004): 542–63.

Pop-Eleches, Grigore. 'Throwing out the Buns: Protest Voting and Unorthodox Parties after Communism'. *World Politics* 62, no. 2 (2010): 221–60.

Rooduijn, Matthijs, Sarah L. De Lange, and Wouter Van Der Brug. 'A Populist Zeitgeist? Programmatic Contagion by Populist Parties in Western Europe'. *Party Politics* 20, no. 4 (2014): 563–75.

Spoon, Jae-Jae, and Heike Klüver. 'Responding to Far Right Challengers: Does Accommodation Pay Off?' *Journal of European Public Policy* 27, no. 2 (1 February 2020): 273–91.

Tankard, Margaret E., and Elizabeth Levy Paluck. "Norm Perception as a Vehicle for Social Change." *Social Issues and Policy Review* 10, no. 1 (2016): 181-211.

Taylor, Sean J., Lev Muchnik, Madhav Kumar, and Sinan Aral. 'Identity Effects in Social Media'. *Nature Human Behaviour*, 10 November 2022, 1–11.

Valentim, Vicente. "Parliamentary Representation and the Normalization of Radical Right Support." *Comparative Political Studies*, 2021.

Valentim, Vicente, and Tobias Widmann. 'Does Radical-Right Success Make the Political Debate More Negative? Evidence from Emotional Rhetoric in German State Parliaments'. *Political Behavior*, 24 March 2021.

Van Spanje, Joost, and Wouter Van Der Brug. "The Party as Pariah: The Exclusion of Anti-Immigration Parties and Its Effect on Their Ideological Positions." *West European Politics* 30, no. 5 (2007): 1022–40.

Ziblatt, Daniel. *Conservative Political Parties and the Birth of Modern Democracy in Europe*. Cambridge University Press, 2017.

Online Appendix

Appendix A: Deviations from the pre-analysis plan

Our pre-analyses plan specified hypotheses for the main effect only. We did specify that we would look into heterogeneous treatment effects based on pre-treatment covariates (like partisanship) but we did not pre-register any hypotheses regarding them. After seeing the heterogeneous effects based on partisanship, we found these patterns extremely interesting, and felt that they would be in line with previous work. In fact, we found these heterogeneous treatment effects more interesting than the main effects. For this reason, we ended up writing the paper around this heterogeneity, including discussing in the theory section how previous work would lead us to expect such heterogeneity. This being said, we want to be clear that, even if the models are as pre-registered, the specific expectations regarding heterogeneity based on partisanship are post-hoc. With this decision, we hope to give adequate weight to a finding that we find very important and in line with previous research, while at the same time not sacrificing transparency and acknowledging that we did not pre-register these specific expectations–only as exploratory additional analyses.

Appendix B: Details About Treatment Vignettes

Table B.1. Sources of the quotes for the different treatment vignettes.

Label	Speaker	Setting
AfD candidate 1 (RRP Approve)	AfD Chair Frauke Petry	Interview with SPIEGEL, published 3/3/2016 https://www.spiegel.de/international/germany/interview-with-frauke-petry-of-the-alternative-for-germany-a-1084493.html
AfD candidate 2 (RRP Approve)	AfD Deputy Leader Beatrix von Storch	Statement by Beatrix von Storch, quoted on 07/11/2019 https://www.infomigrants.net/en/post/17960/a-plan-to-stop-immigration-germanys-afd-party
CDU candidate 1 (Mainstream Approve)	CDU leadership candidate Friedrich Merz	Comment by Friedrich Merz in Seebach, quoted on 22/11/2018 https://www.ft.com/content/99c8f214-ee31-11e8-89c8-d36339d835c0
CDU candidate 2 (Mainstream Approve)	CDU premier of Hesse, Roland Koch	Interview with *Bild*, quoted on 01/09/2008 https://www.dw.com/en/german-ex-chancellor-criticizes-xenophobic-election-campaign/a-3048492
CDU candidate (Mainstream Disapprove)	CDU Chancellor Angela Merkel	Speech on the 70th anniversary of the passing of the German Constitution, quoted 15/5/2019 https://www.thelocal.de/20190515/germanys-future-depends-on-immigration-and-integration-merkel/
Candidate (Control)	CDU Federal Minister of Food, Agriculture and Consumer Protection	Article written for the European Commission on food supply chains, published 10/08/2020 Initially available at: https://ec.europa.eu/cyprus/news/20201008_5_en+&cd=1&hl=en&ct=clnk&gl=us

Table B.2. Coding scores for the different treatment vignettes.

	Please could you score the statements on a scale of 1-9, with 1 indicating that Party A's statements are much more xenophobic, and 9 indicating that Party B's statements are much more xenophobic, and 5 indicating that the statements from both Parties are equally xenophobic?	**Able to identify Party A speakers**	**Able to identify Party B speakers**
Coder 1	5	No	No
Coder 2	5	No	No
Coder 3	6	No	No
Coder 4	5	No	No

Note: Party A quotes were those from the RRP Approve *condition*, and *Party B quotes were those from the* Mainstream Approve *condition*.

Appendix C: Details on the experimental treatment

Table C.1: Vignettes that are presented to survey respondents based on the condition they are cross-randomized into.

RRP Approve
We will now show you two actual statements made by candidates from the **AfD.**
AfD candidate 1: "Ms. Merkel simply opened the borders and invited everybody in…Such a debate [about immigration] is imperative because the economic and social consequences on both home and host countries are equally momentous.
AfD candidate 2: "There have been rising numbers of criminal rates within those migrant groups."
Mainstream Approve
We will now show you two actual statements made by candidates from the **CDU**.
CDU candidate 1: "I have long thought that we should be prepared to openly discuss this constitutional right to asylum and whether it can continue in this form if we seriously want a European immigration and refugee policy."
CDU candidate 2: "How much are we prepared to take from a small proportion of violent youths, who frequently have a foreign background?"
Mainstream Approve + RRP Approve
We will now show you four actual statements. The first and second were made by candidates from the **CDU**, and the third and fourth were made by candidates from the **AfD**.
CDU candidate 1: "I have long thought that we should be prepared to openly discuss this constitutional right to asylum and whether it can continue in this form if we seriously want a European immigration and refugee policy."
CDU candidate 2: "How much are we prepared to take from a small proportion of violent youths, who frequently have a foreign background?"
AfD candidate 1: "Ms. Merkel simply opened the borders and invited everybody in…Such a debate [about immigration] is imperative because the economic and social consequences on both home and host countries are equally momentous.
AfD candidate 2: "There have been rising numbers of criminal rates within those migrant groups."
Mainstream Disapprove + RRP Approve
We will now show you three actual statements. The first was made by a candidate from the **CDU**, and the second and third were made by candidates from the **AfD**.
CDU candidate: "We have learned that our country must be a country of immigration and of integration."

AfD candidate 1: "Ms. Merkel simply opened the borders and invited everybody in…Such a debate [about immigration] is imperative because the economic and social consequences on both home and host countries are equally momentous.

AfD candidate 2: "There have been rising numbers of criminal rates within those migrant groups."

Control

We will now show you an actual statement made by a candidate from a political party.

Candidate: "We can develop new business models across the entire food supply chain, allowing us to avoid overproduction whilst using valuable resources more efficiently."

Figure C.1: The hypothetical petition that participants are presented.

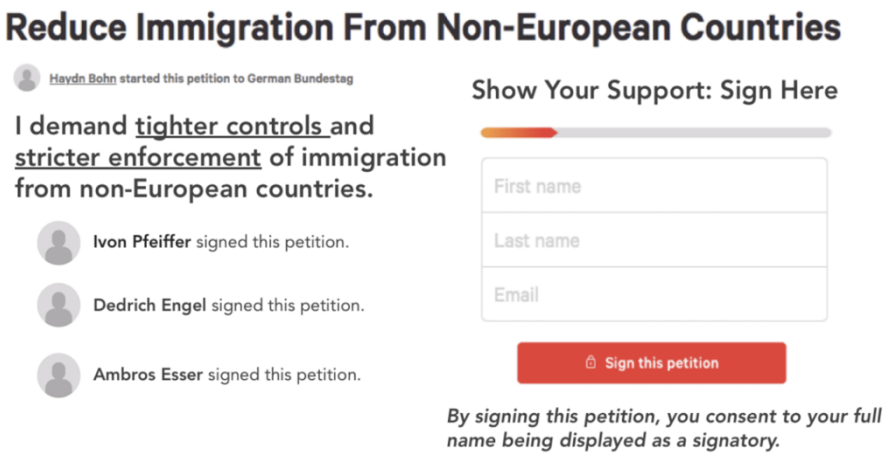


Figure C.2: Example tweet that respondent is shown.

The user was suspended from platform after violating Twitter's hateful conduct policy.

Figure C.3: Mean values for personal views about appropriateness (95% CIs).

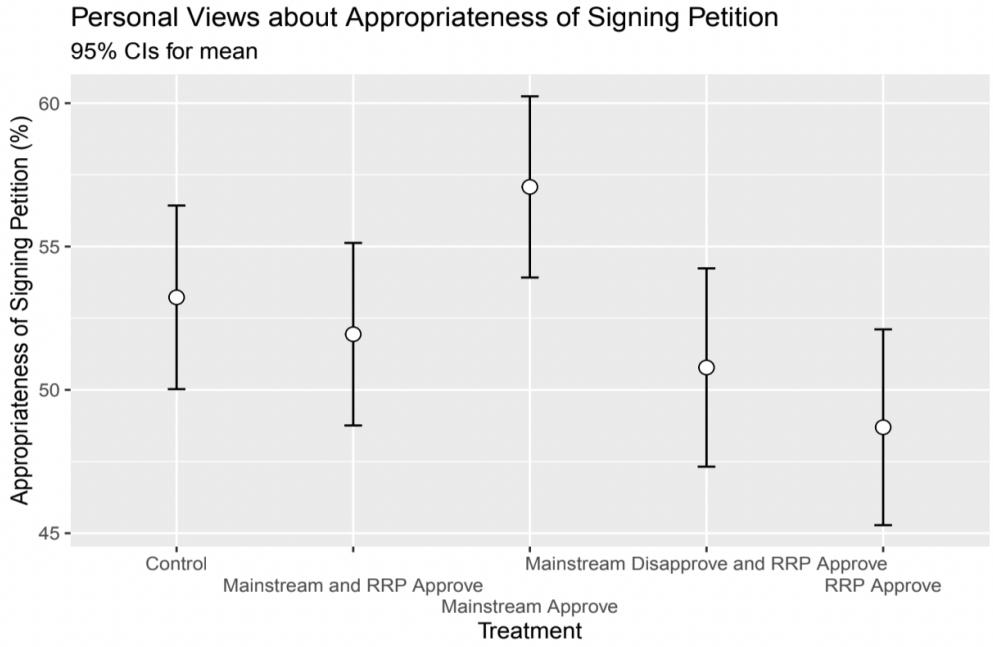
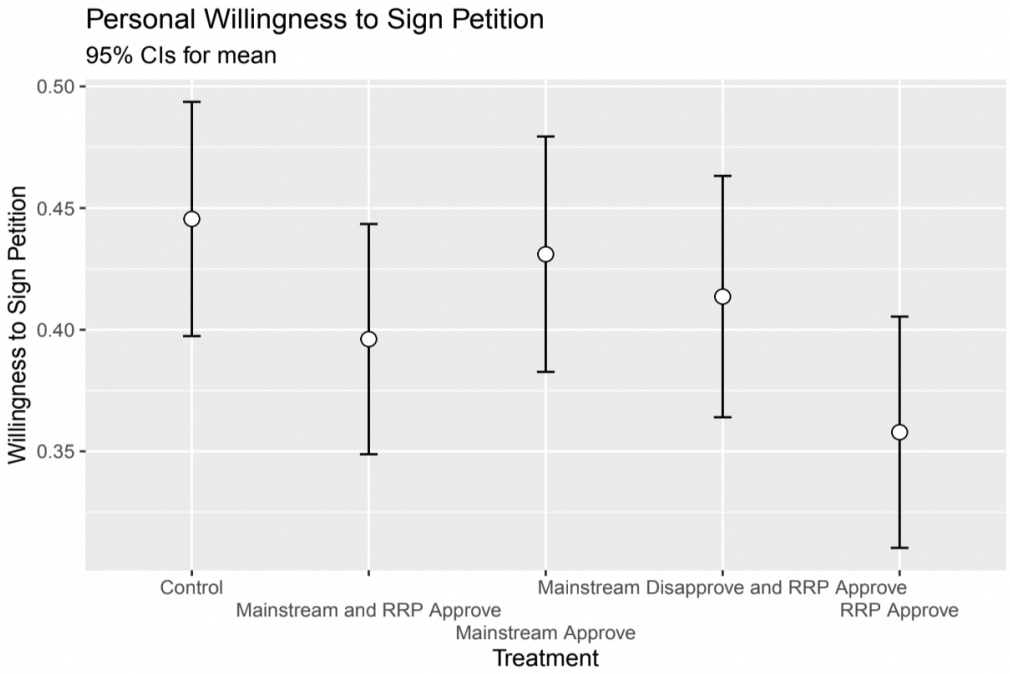


Figure C.4: Mean values for personal willingness to sign (95% CIs).



Appendix D: Differential Attrition Check

Table D.1: Number of screenouts by treatment condition, compared against the total number of respondents that were assigned to each treatment arm.

Treatment	Count in screenouts	Count in total sample	% of total sample
Control	39	452	8.63
Mainstream and RRP Approve	41	455	9.01
Mainstream Approve	38	444	8.56
Mainstream Disapprove and RRP Approve	38	420	9.05
RRP Approve	43	437	9.84

Appendix E: Full Results of the Main Analyses

Full Results of Figure 1.

Table E.1: Agreement with Sensitive Item.

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	0.06	0.07	0.05	0.00
	[-0.04, 0.15]	[-0.11, 0.25]	[-0.09, 0.19]	[-0.37, 0.37]
MRPapprovedummy	-0.01	-0.06	0.03	0.15
	[-0.15, 0.13]	[-0.32, 0.20]	[-0.17, 0.24]	[-0.35, 0.64]
RRPapprovedummy	-0.12	-0.05	-0.20 *	-0.15
	[-0.25, 0.02]	[-0.31, 0.21]	[-0.39, -0.01]	[-0.73, 0.42]
MRPdisapproveRRPapprovedummy	-0.01	-0.08	0.04	-0.14
	[-0.15, 0.13]	[-0.35, 0.19]	[-0.16, 0.24]	[-0.68, 0.40]
MRPapproveRRPapprovedummy	-0.14 *	-0.19	-0.13	0.25
	[-0.28, -0.01]	[-0.44, 0.07]	[-0.32, 0.06]	[-0.21, 0.71]
N	2009	545	1021	135
R2	0.00	0.00	0.01	0.02

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.2: Personal Willingness to Sign.

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	0.07	0.08	0.06	0.05
	[-0.02, 0.17]	[-0.10, 0.26]	[-0.08, 0.20]	[-0.28, 0.39]
MRPapprovedummy	-0.03	-0.02	0.06	-0.21
	[-0.17, 0.11]	[-0.28, 0.24]	[-0.14, 0.26]	[-0.81, 0.39]
RRPapprovedummy	-0.18 *	-0.09	-0.24 *	-0.16
	[-0.32, -0.04]	[-0.35, 0.17]	[-0.42, -0.05]	[-0.72, 0.39]
MRPdisapproveRRPapprovedummy	-0.06	-0.15	-0.08	0.13
	[-0.21, 0.08]	[-0.42, 0.12]	[-0.28, 0.11]	[-0.27, 0.53]
MRPapproveRRPapprovedummy	-0.10	-0.15	-0.03	-0.10
	[-0.24, 0.04]	[-0.40, 0.10]	[-0.23, 0.17]	[-0.68, 0.47]
N	2009	545	1021	135
R2	0.00	0.00	0.01	0.02

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.3: Personal Views of Appropriateness of Signing

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	0.03	-0.05	0.06	-0.21
	[-0.07, 0.12]	[-0.23, 0.12]	[-0.08, 0.20]	[-0.64, 0.22]
MRPapprovedummy	0.11	0.22	0.12	0.44
	[-0.02, 0.25]	[-0.03, 0.47]	[-0.08, 0.31]	[-0.07, 0.95]
RRPapprovedummy	-0.14	0.06	-0.27 **	0.07
	[-0.27, 0.00]	[-0.20, 0.33]	[-0.46, -0.08]	[-0.59, 0.74]
MRPdisapproveRRPapprovedummy	-0.07	-0.05	-0.11	0.16
	[-0.21, 0.07]	[-0.33, 0.22]	[-0.31, 0.09]	[-0.37, 0.69]
MRPapproveRRPapprovedummy	-0.04	0.04	-0.06	0.50 *
	[-0.17, 0.10]	[-0.22, 0.29]	[-0.25, 0.14]	[0.03, 0.97]
N	2009	545	1021	135
R2	0.01	0.01	0.02	0.04

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.4: Empirical Expectations

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.04	-0.14	0.07	0.01
	[-0.14, 0.06]	[-0.32, 0.05]	[-0.07, 0.21]	[-0.33, 0.34]
MRPapprovedummy	0.15 *	0.31 *	0.04	0.19
	[0.02, 0.29]	[0.05, 0.57]	[-0.15, 0.24]	[-0.39, 0.77]
RRPapprovedummy	-0.06	0.12	-0.26 **	-0.06
	[-0.20, 0.08]	[-0.15, 0.39]	[-0.46, -0.06]	[-0.58, 0.45]
MRPdisapproveRRPapprovedummy	0.01	-0.05	-0.05	-0.19
	[-0.13, 0.15]	[-0.33, 0.23]	[-0.25, 0.14]	[-0.65, 0.28]
MRPapproveRRPapprovedummy	0.08	0.28 *	-0.09	0.11
	[-0.06, 0.22]	[0.03, 0.53]	[-0.29, 0.10]	[-0.42, 0.65]
N	2009	545	1021	135
R2	0.01	0.02	0.01	0.02

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.5: Normative Expectations

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	0.01	-0.11	0.07	0.26
	[-0.08, 0.11]	[-0.30, 0.08]	[-0.06, 0.21]	[-0.05, 0.57]
MRPapprovedummy	0.12	0.31 *	0.07	-0.09
	[-0.01, 0.26]	[0.05, 0.57]	[-0.12, 0.25]	[-0.69, 0.51]
RRPapprovedummy	-0.08	0.09	-0.23 *	-0.26
	[-0.22, 0.06]	[-0.19, 0.36]	[-0.43, -0.03]	[-0.74, 0.22]
MRPdisapproveRRPapprovedummy	-0.09	-0.14	-0.07	-0.58 *
	[-0.23, 0.05]	[-0.41, 0.14]	[-0.26, 0.12]	[-1.04, -0.12]
MRPapproveRRPapprovedummy	-0.02	0.24	-0.14	-0.33
	[-0.16, 0.11]	[-0.01, 0.50]	[-0.33, 0.05]	[-0.81, 0.15]
N	2009	545	1021	135
R2	0.01	0.03	0.01	0.05

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.6: Sanctioning

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.00	-0.04	-0.02	0.33
	[-0.10, 0.09]	[-0.22, 0.14]	[-0.16, 0.11]	[-0.12, 0.79]
MRPapprovedummy	0.04	0.02	0.10	-0.40
	[-0.10, 0.18]	[-0.24, 0.28]	[-0.09, 0.29]	[-0.99, 0.20]
RRPapprovedummy	-0.02	0.09	0.02	-0.64 *
	[-0.15, 0.12]	[-0.17, 0.36]	[-0.17, 0.22]	[-1.14, -0.14]
MRPdisapproveRRPapprovedummy	0.02	0.09	0.01	-0.24
	[-0.12, 0.16]	[-0.17, 0.36]	[-0.19, 0.20]	[-0.83, 0.36]
MRPapproveRRPapprovedummy	-0.03	0.01	-0.01	-0.48
	[-0.17, 0.10]	[-0.25, 0.26]	[-0.21, 0.18]	[-1.06, 0.10]
N	2009	545	1021	135
R2	0.00	0.00	0.00	0.05

Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Replication of Figure 2.

Table E.7: Agreement with Sensitive Item

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.05	0.00	-0.12	-0.14
	[-0.15, 0.04]	[-0.19, 0.20]	[-0.25, 0.00]	[-0.58, 0.30]
MRPapprovedummy	0.11	-0.01	0.24 *	0.29
	[-0.03, 0.24]	[-0.28, 0.26]	[0.05, 0.44]	[-0.26, 0.85]
N	800	211	404	51
R2	0.00	0.00	0.01	0.02

Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table E.7: Personal Willingness to Sign Petition

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.08	-0.03	-0.15 *	0.02
	[-0.17, 0.02]	[-0.23, 0.16]	[-0.28, -0.03]	[-0.36, 0.40]
MRPapprovedummy	0.15 *	0.06	0.30 **	-0.04
	[0.01, 0.29]	[-0.21, 0.34]	[0.11, 0.49]	[-0.61, 0.53]
N	800	211	404	51
R2	0.01	0.00	0.02	0.00

Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table E.8: Views on Appropriateness of Signing

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.13 *	-0.08	-0.20 **	-0.16
	[-0.23, -0.03]	[-0.28, 0.12]	[-0.33, -0.07]	[-0.63, 0.32]
MRPapprovedummy	0.25 ***	0.16	0.40 ***	0.34
	[0.11, 0.39]	[-0.11, 0.43]	[0.20, 0.59]	[-0.21, 0.88]
N	800	211	404	51
R2	0.02	0.01	0.04	0.03

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.9: Empirical Expectations

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.11 *	-0.10	-0.16 *	-0.11
	[-0.21, -0.01]	[-0.30, 0.10]	[-0.30, -0.02]	[-0.47, 0.25]
MRPapprovedummy	0.22 **	0.19	0.31 **	0.23
	[0.08, 0.36]	[-0.08, 0.47]	[0.11, 0.50]	[-0.34, 0.80]
N	800	211	404	51
R2	0.01	0.01	0.02	0.01

Standard errors are heteroskedasticity robust. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table E.10: Normative Expectations

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.10 *	-0.12	-0.15 *	-0.07
	[-0.20, -0.00]	[-0.32, 0.08]	[-0.30, -0.00]	[-0.41, 0.26]
MRPapprovedummy	0.20 **	0.23	0.30 **	0.16
	[0.06, 0.34]	[-0.04, 0.50]	[0.10, 0.49]	[-0.42, 0.74]
N	800	211	404	51
R2	0.01	0.01	0.02	0.01

Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Table E.11: Sanctioning

	Full Sample	Right Wing Only	Left Wing Only	Radical Right Only
(Intercept)	-0.03	0.04	-0.04	-0.15
	[-0.13, 0.07]	[-0.15, 0.23]	[-0.18, 0.10]	[-0.43, 0.12]
MRPapprovedummy	0.06	-0.07	0.07	0.32
	[-0.08, 0.20]	[-0.35, 0.20]	[-0.12, 0.27]	[-0.25, 0.90]
N	800	211	404	51
R2	0.00	0.00	0.00	0.03

Standard errors are heteroskedasticity robust. *** p < 0.001; ** p < 0.01; * p < 0.05.

Appendix F: Robustness Checks

Robustness Checks: Adding Controls

Figure F.1: Replication of Figure 1 adding controls.

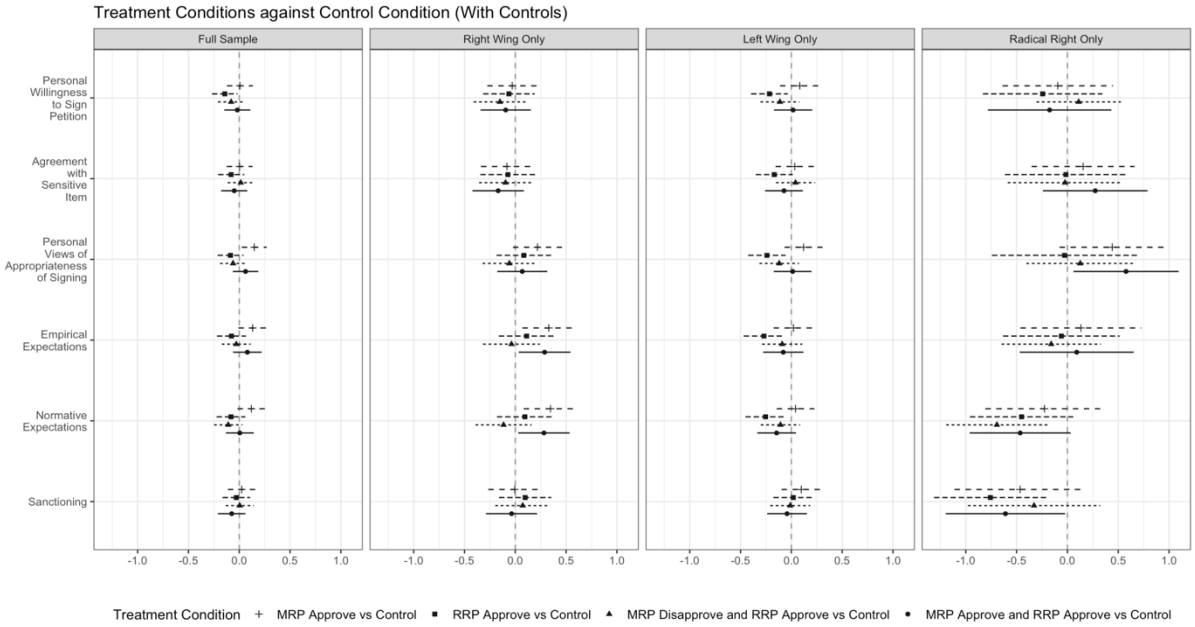
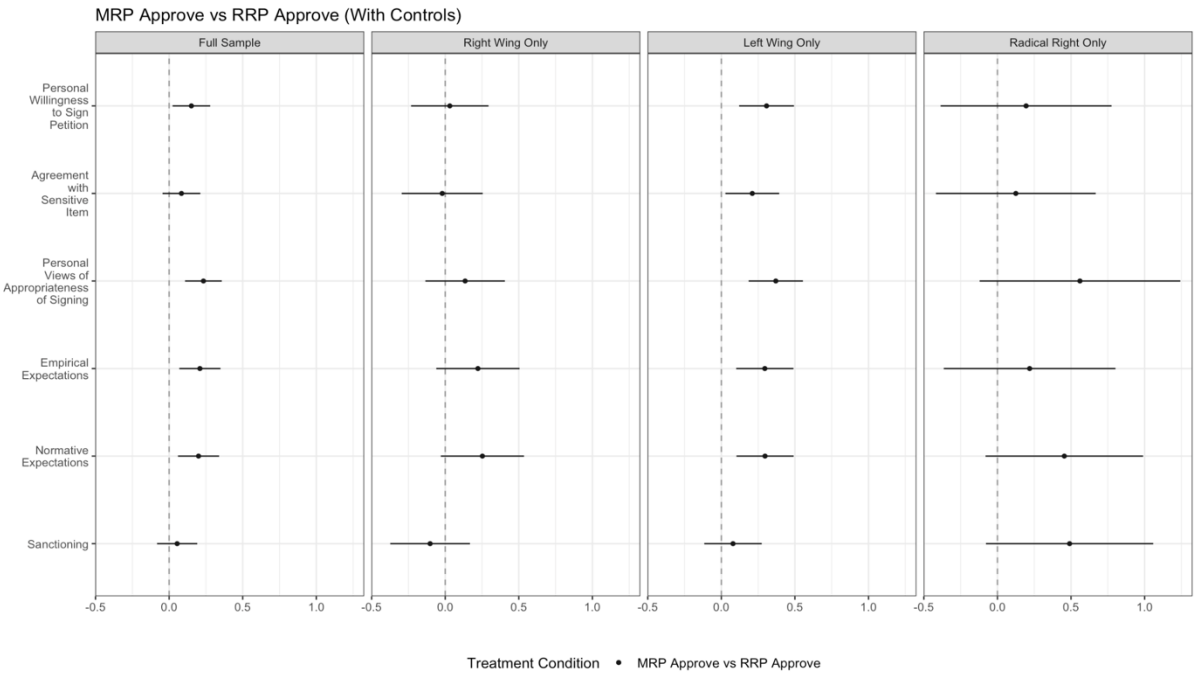


Figure F.2: Replication of Figure 2 adding controls



Heterogenous Effects: Split by Gender

Figure F.3: Comparison of treatment and control conditions with female-only subsample

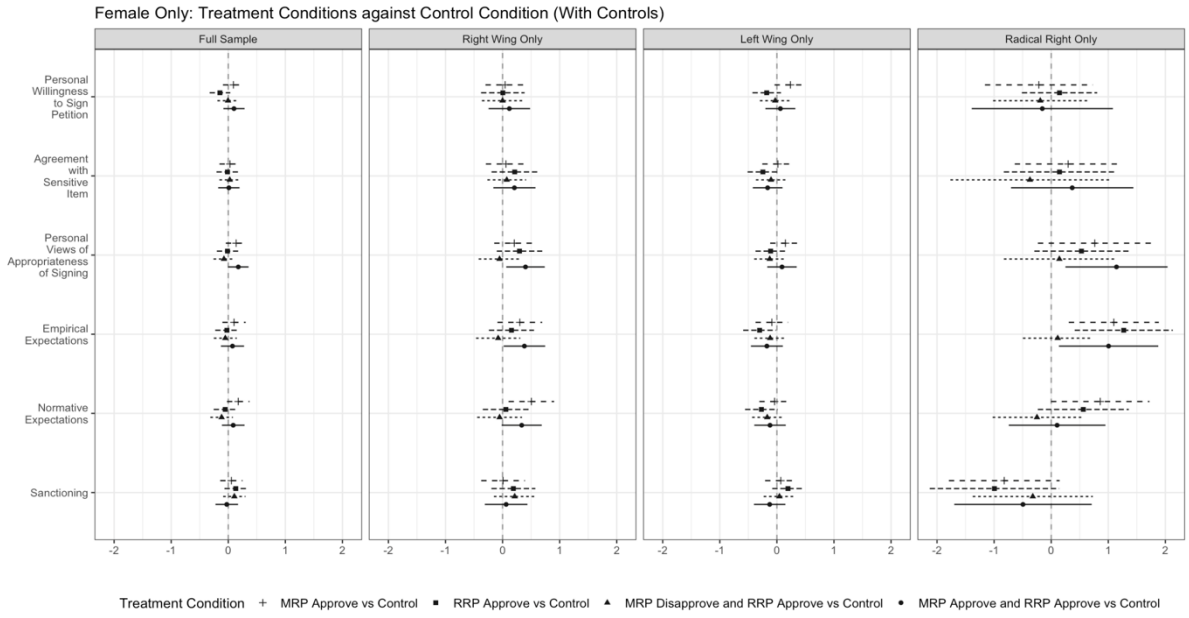
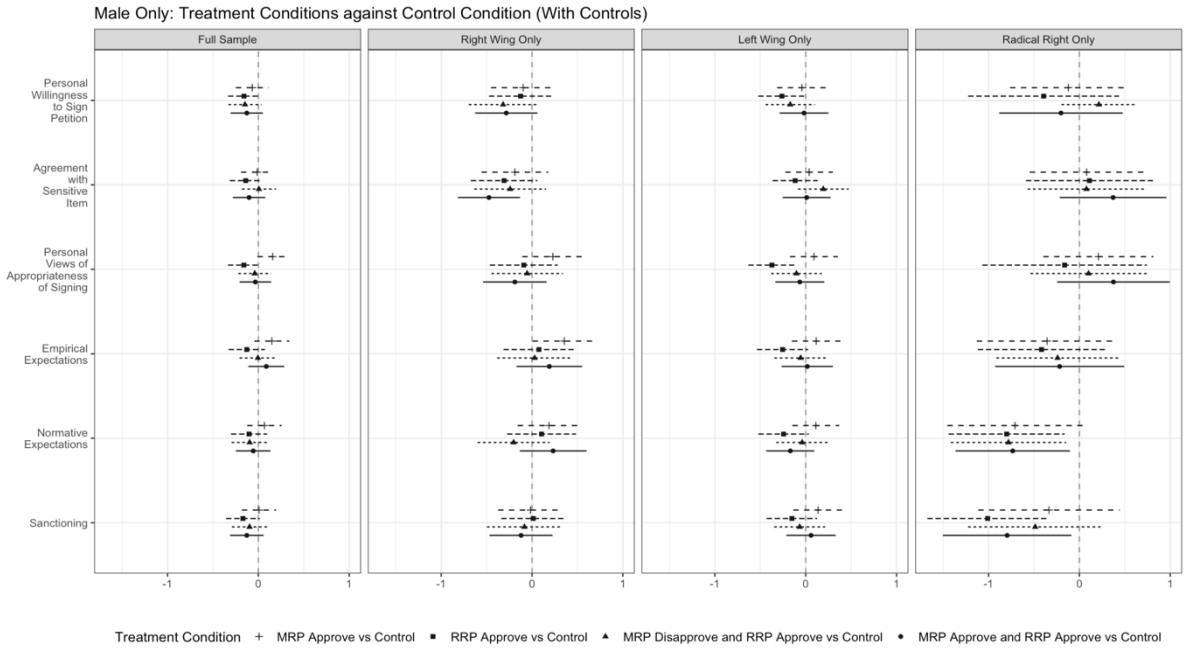


Figure F.4: Comparison of treatment and control conditions with male-only subsample



Heterogenous Effects: Split by Income

Figure F.5: Comparison of treatment and control conditions with subsample who "live comfortably" only

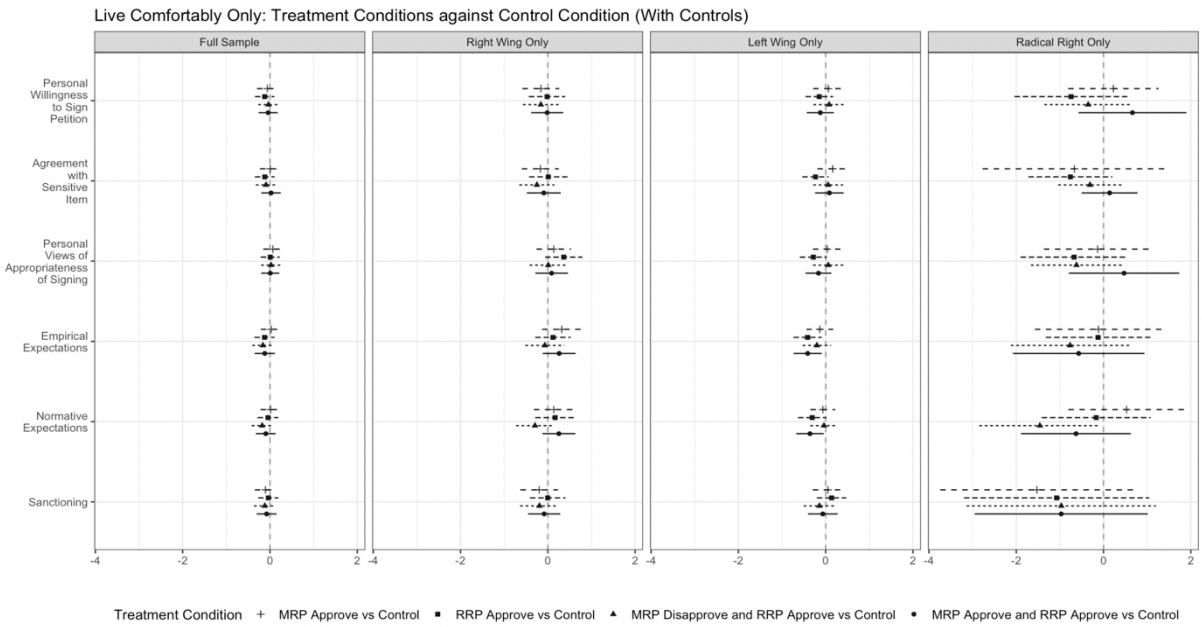


Figure F.6: Comparison of treatment and control conditions with subsample who "make ends meet" only

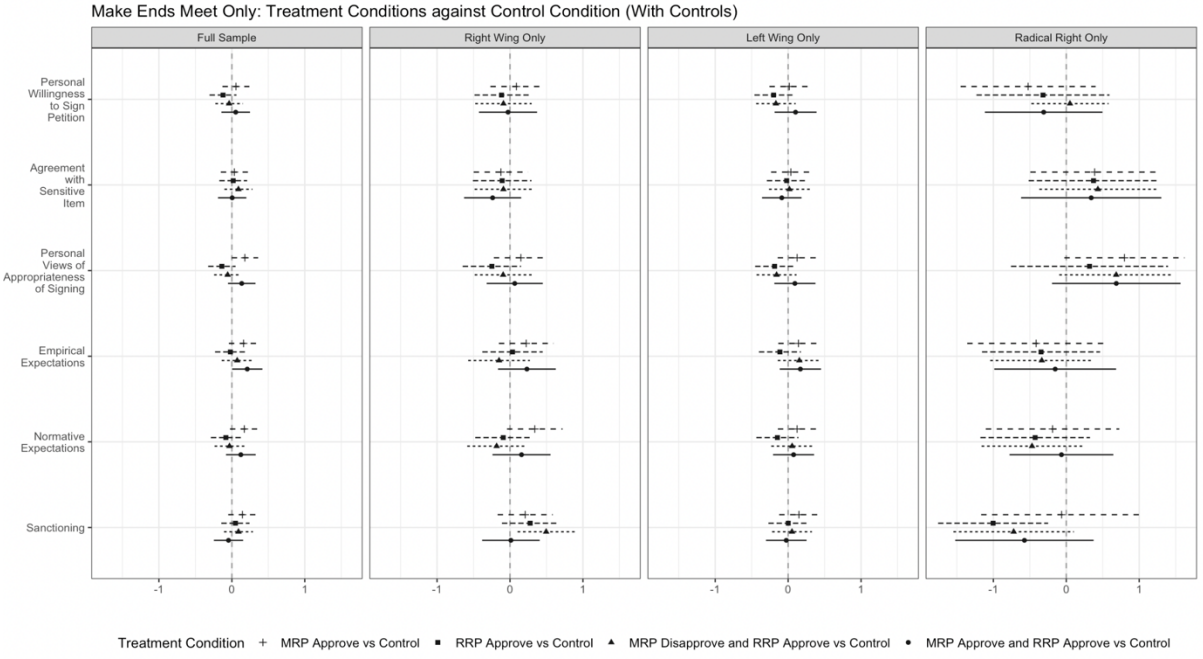
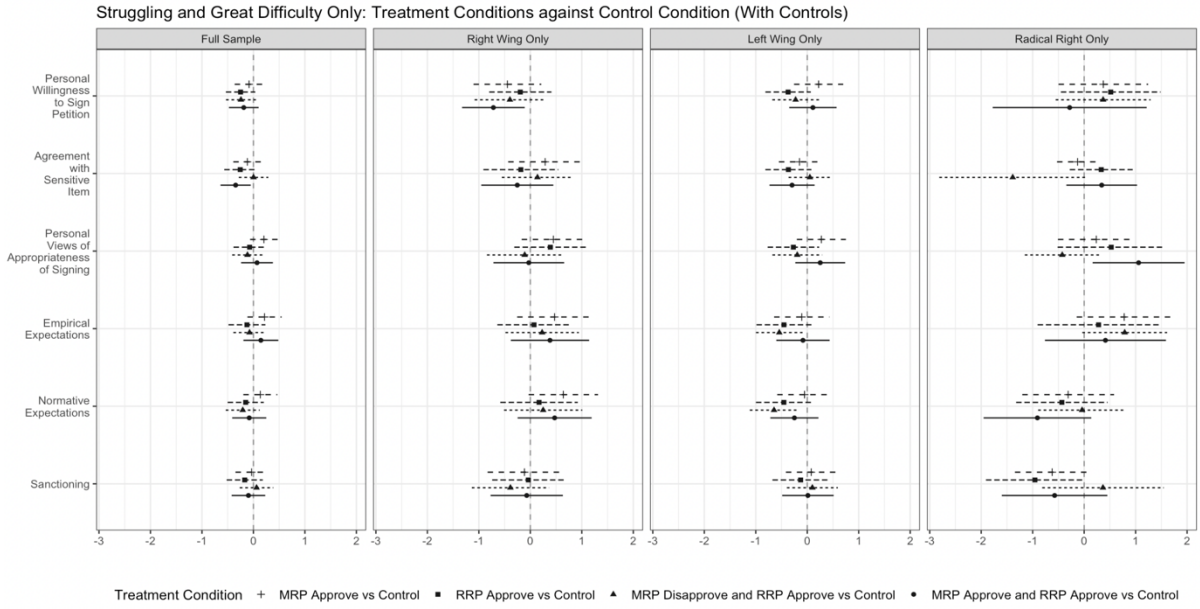


Figure F.7: Comparison of treatment and control conditions with subsample who are "struggling" or "facing great difficulty" only



Heterogenous Effects: Split by Education

Figure F.8: Comparison of treatment and control conditions with subsample with elementary and secondary school educational qualifications

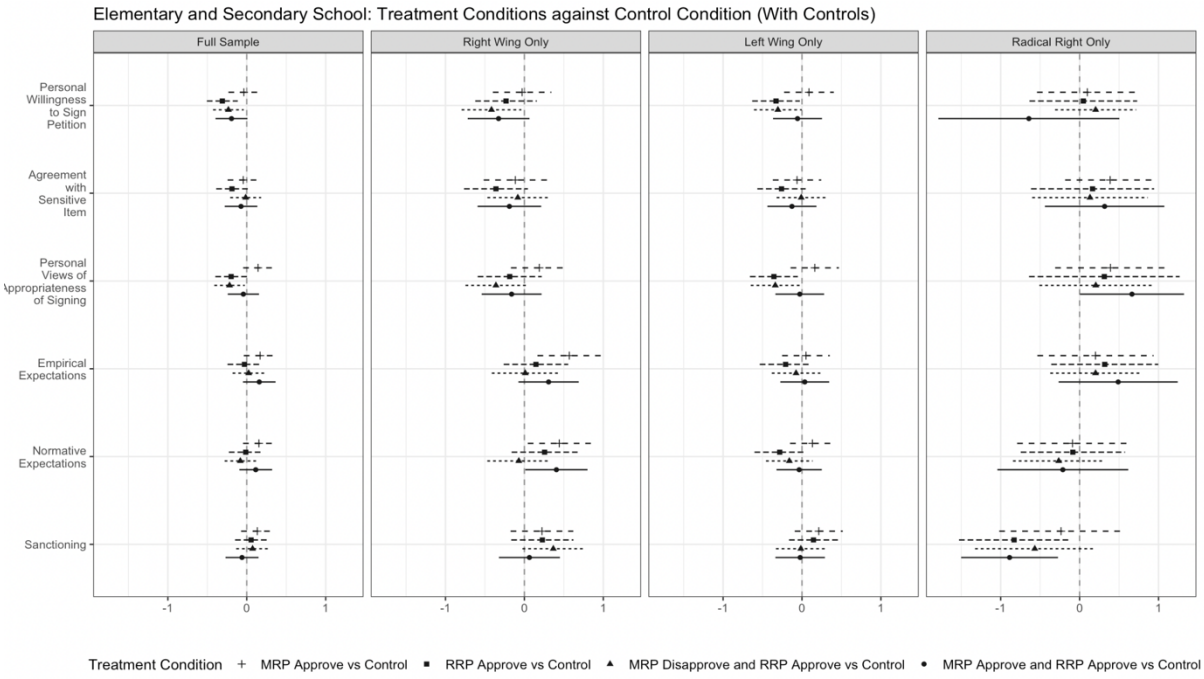


Figure F.9: Comparison of treatment and control conditions with subsample with secondary school and higher education entrance exam educational qualifications

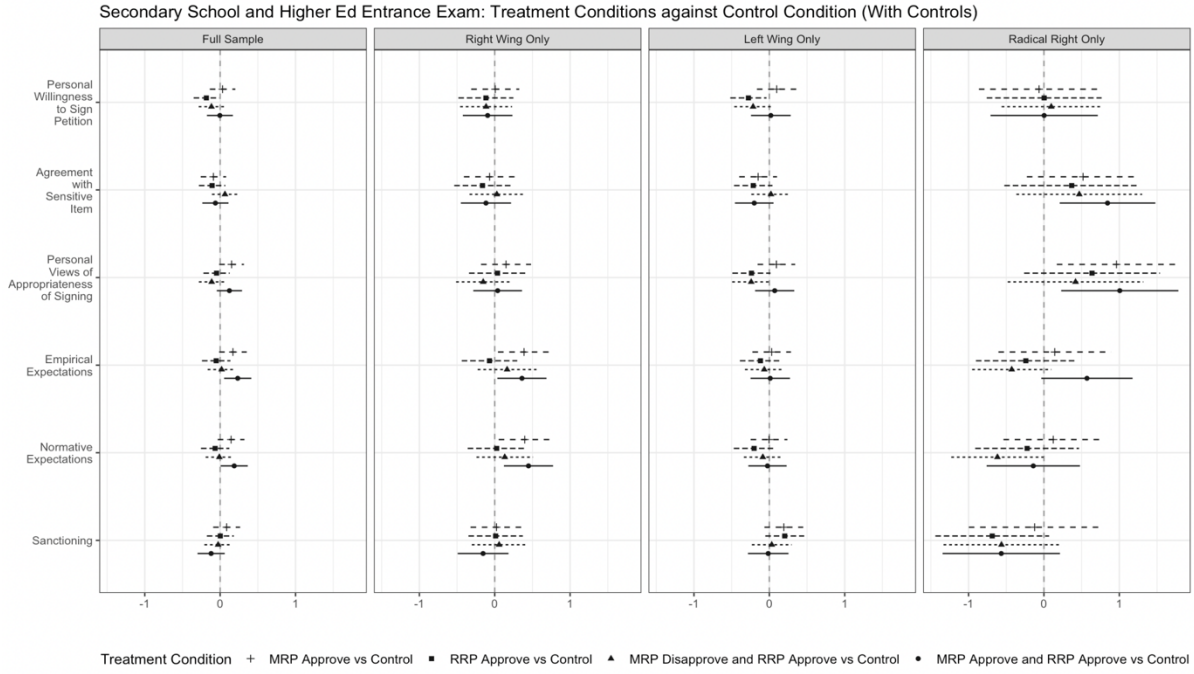


Figure F.10: Comparison of treatment and control conditions with subsample with higher education entrance exam and university educational qualifications

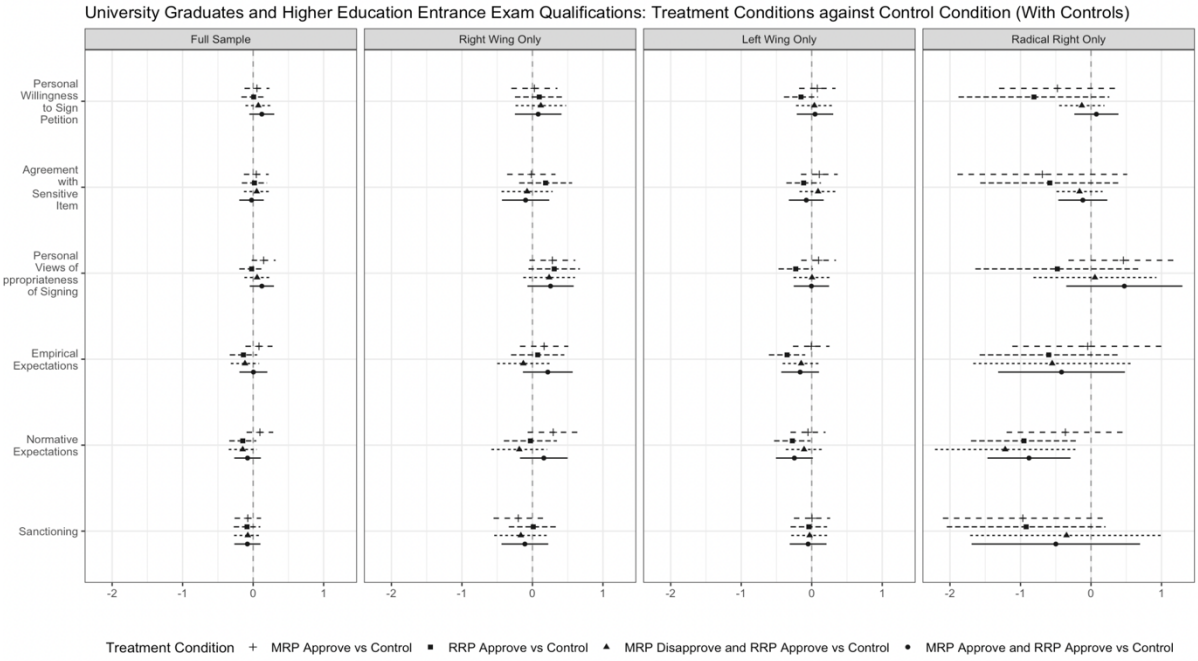
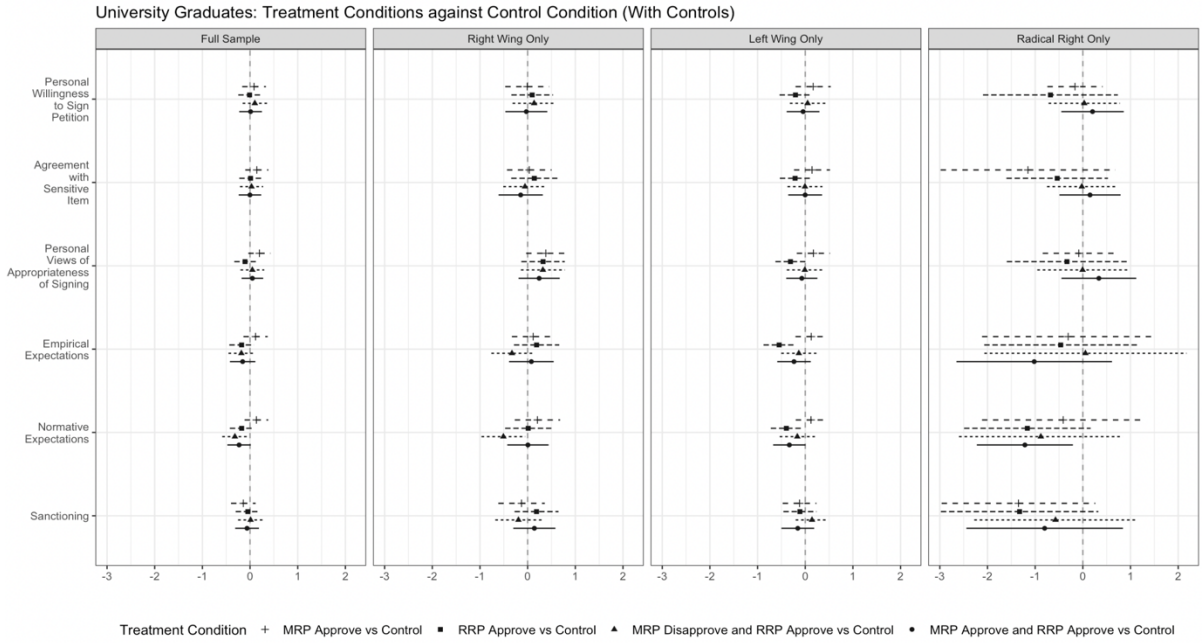


Figure F.11: Comparison of treatment and control conditions with subsample with university graduates only



Sensitivity Analysis: Removing each Control Variable

Figure F.12: Comparison of treatment and control conditions with full controls except age

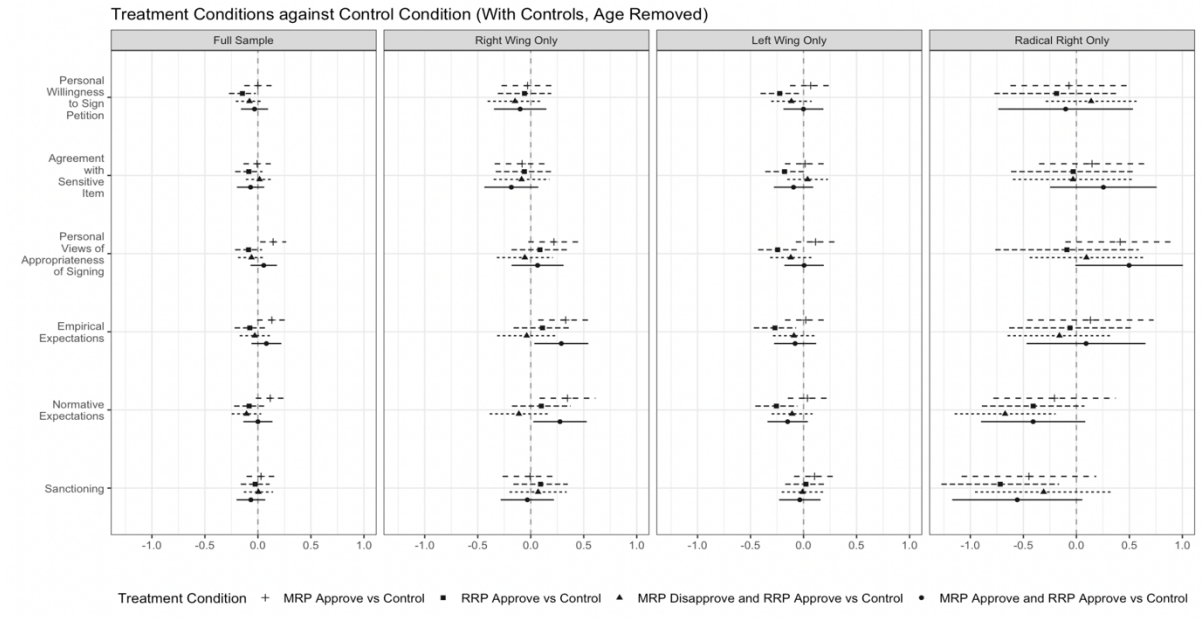


Figure F.13: Comparison of treatment and control conditions with full controls except education

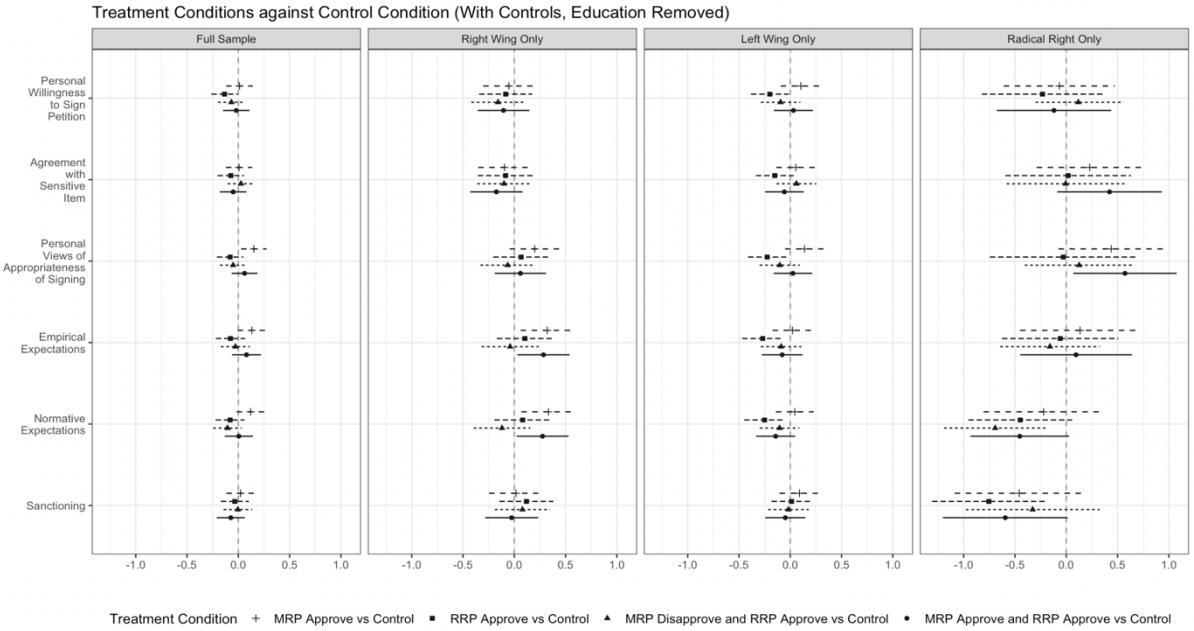


Figure F.14: Comparison of treatment and control conditions with full controls except income

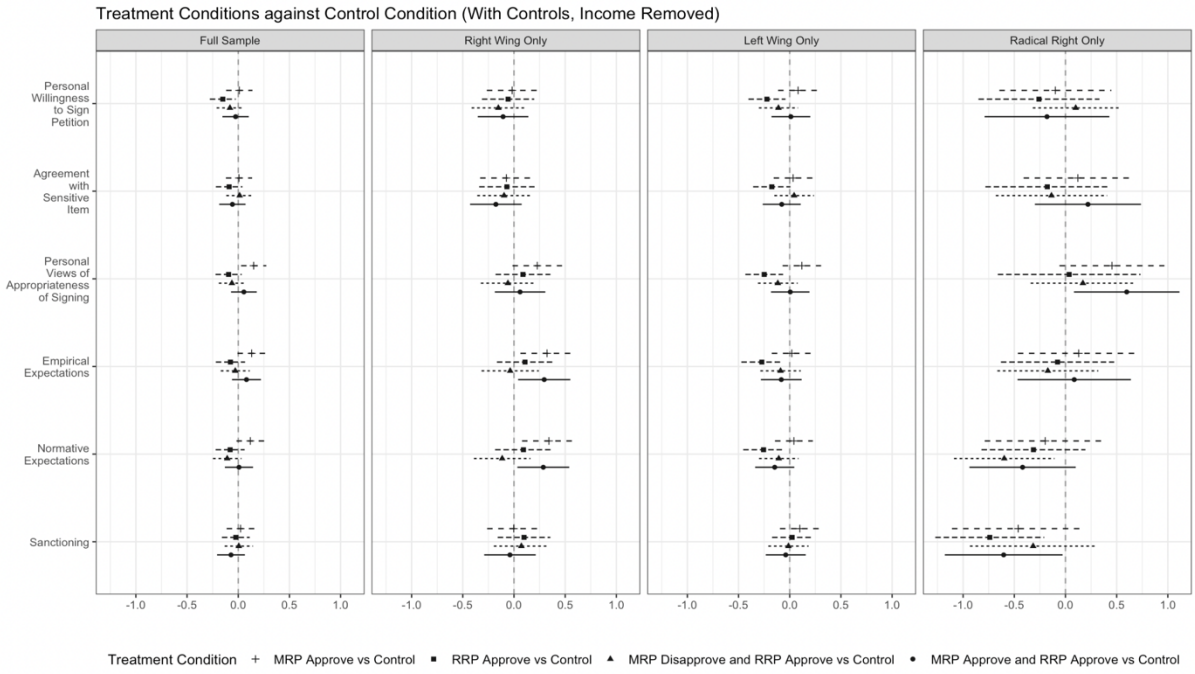
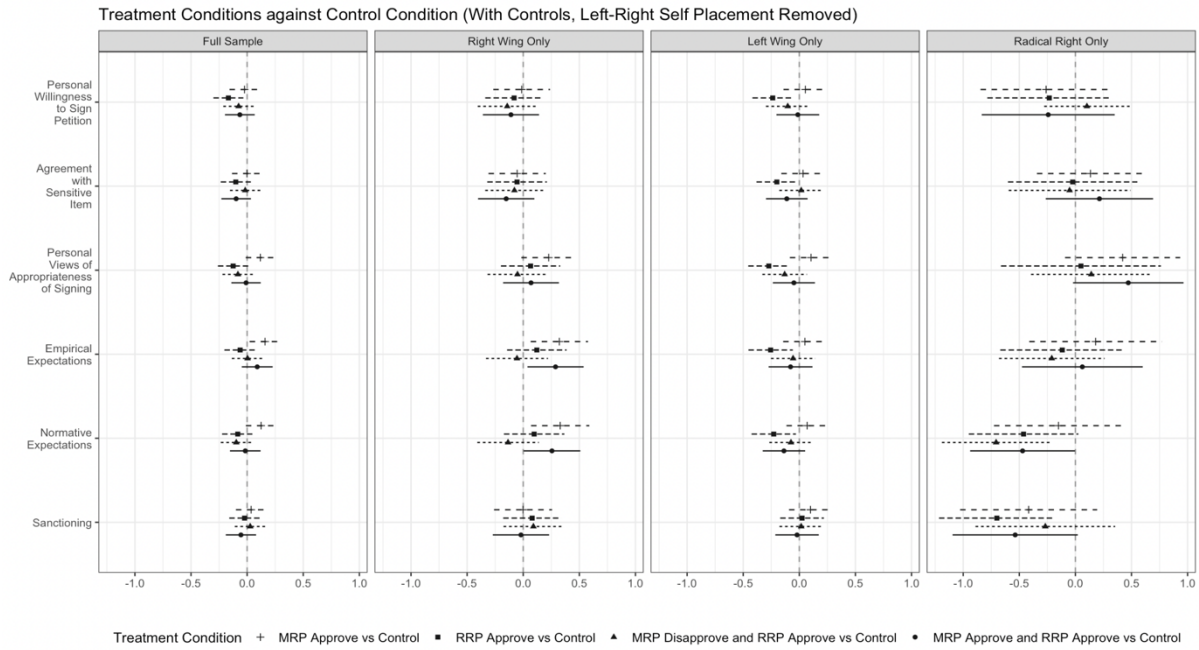


Figure F.15: Comparison of treatment and control conditions with full controls except left-right self-placement



Sensitivity Analysis: Removing each State

Figure F.16: Comparison of treatment and control conditions with all states except Baden-Wuerttemberg

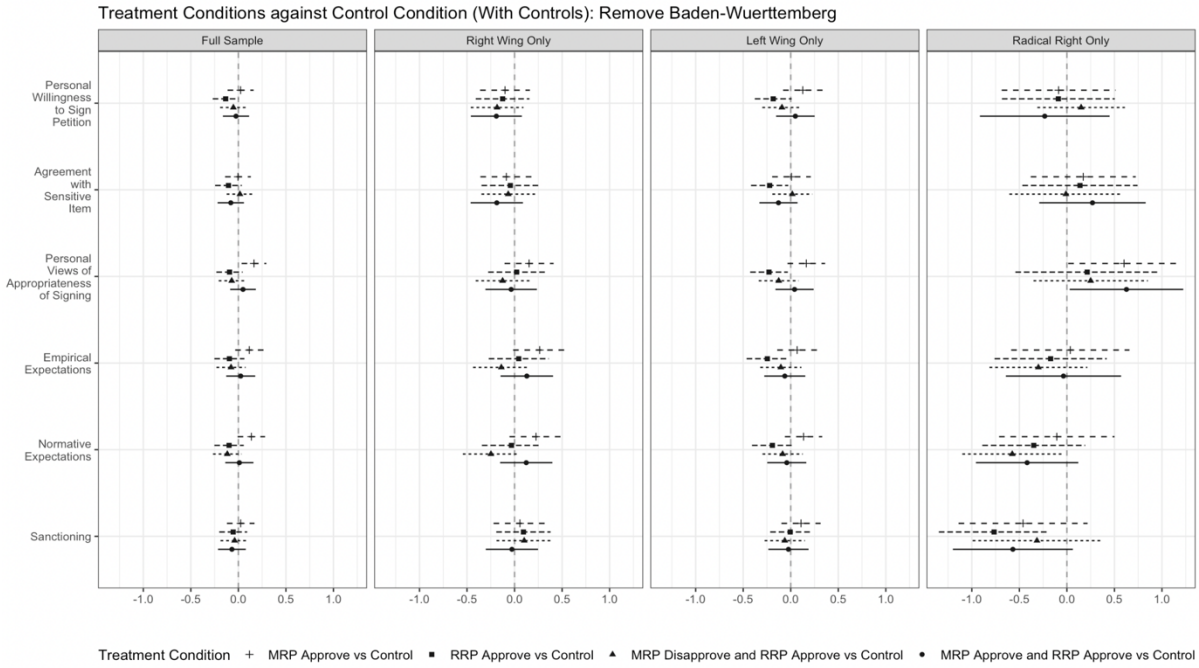


Figure F.17: Comparison of treatment and control conditions with all states except Bayern

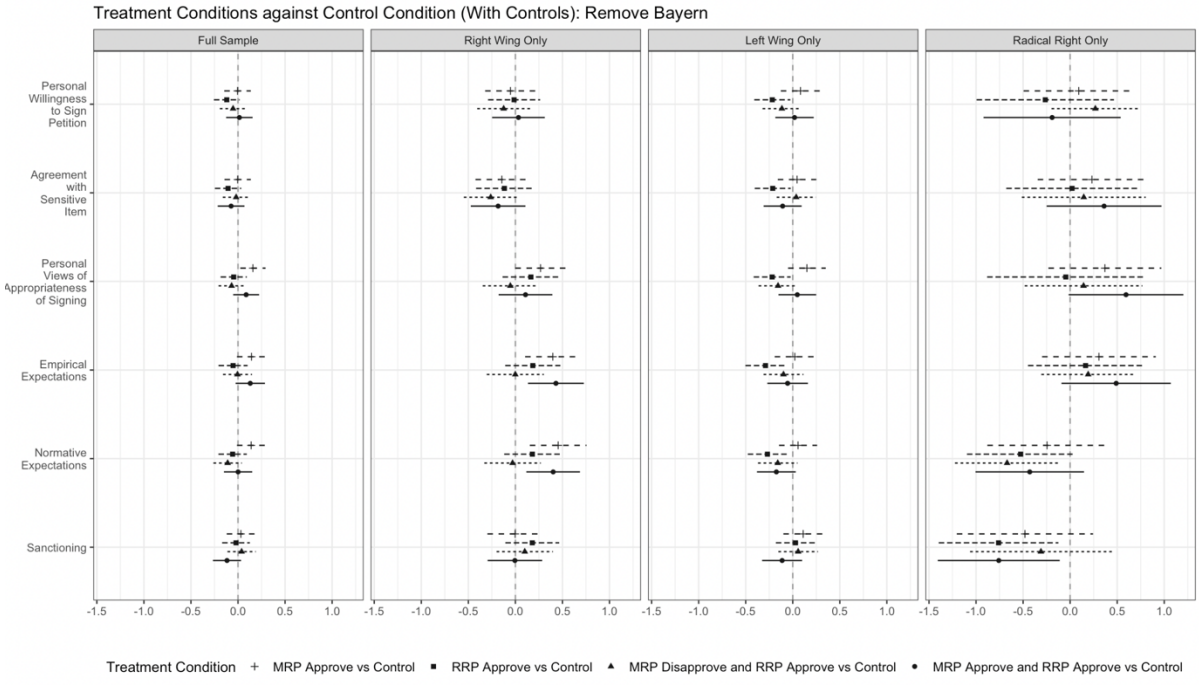


Figure F.18: Comparison of treatment and control conditions with all states except Berlin

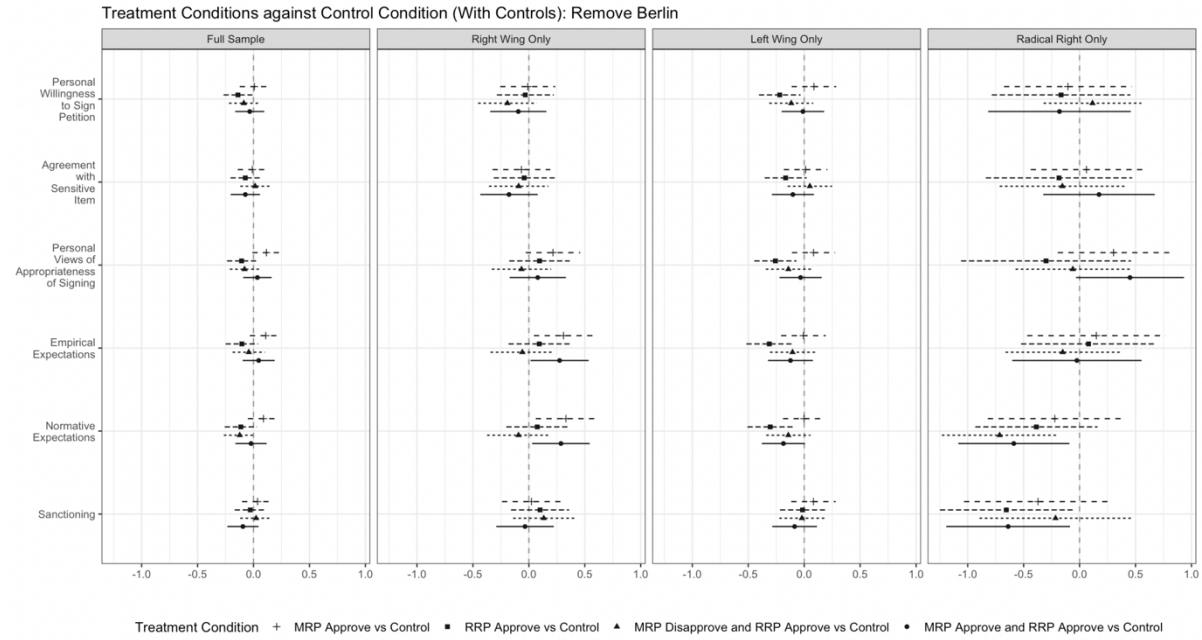


Figure F.19: Comparison of treatment and control conditions with all states except Brandenburg

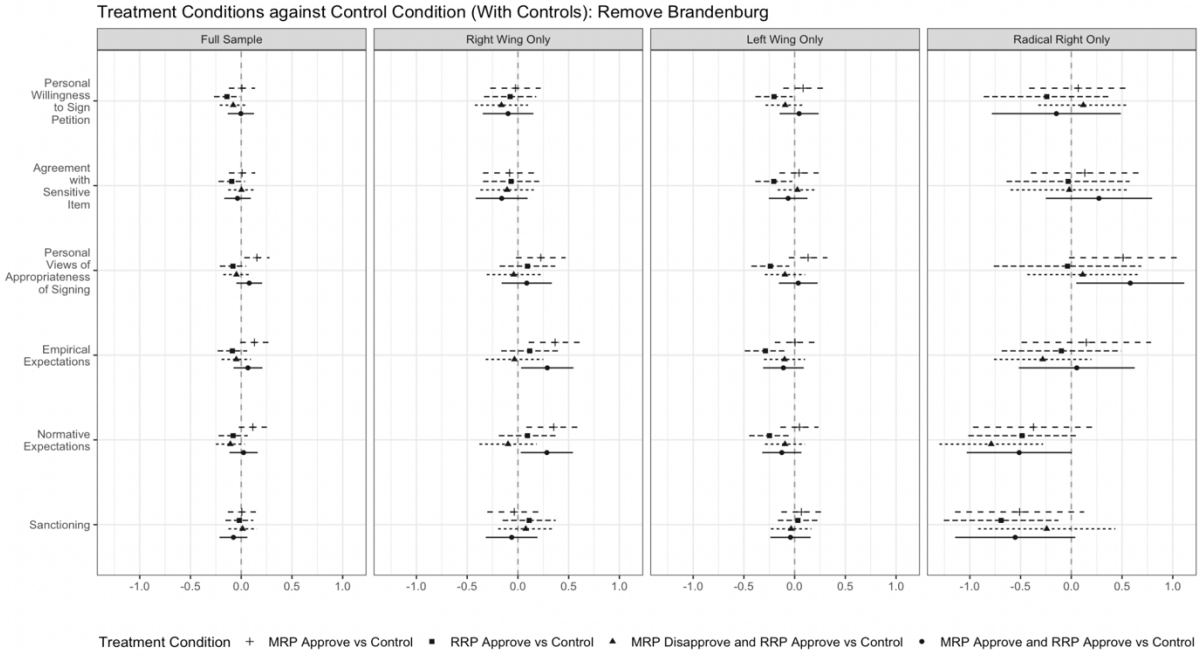


Figure F.20: Comparison of treatment and control conditions with all states except Bremen

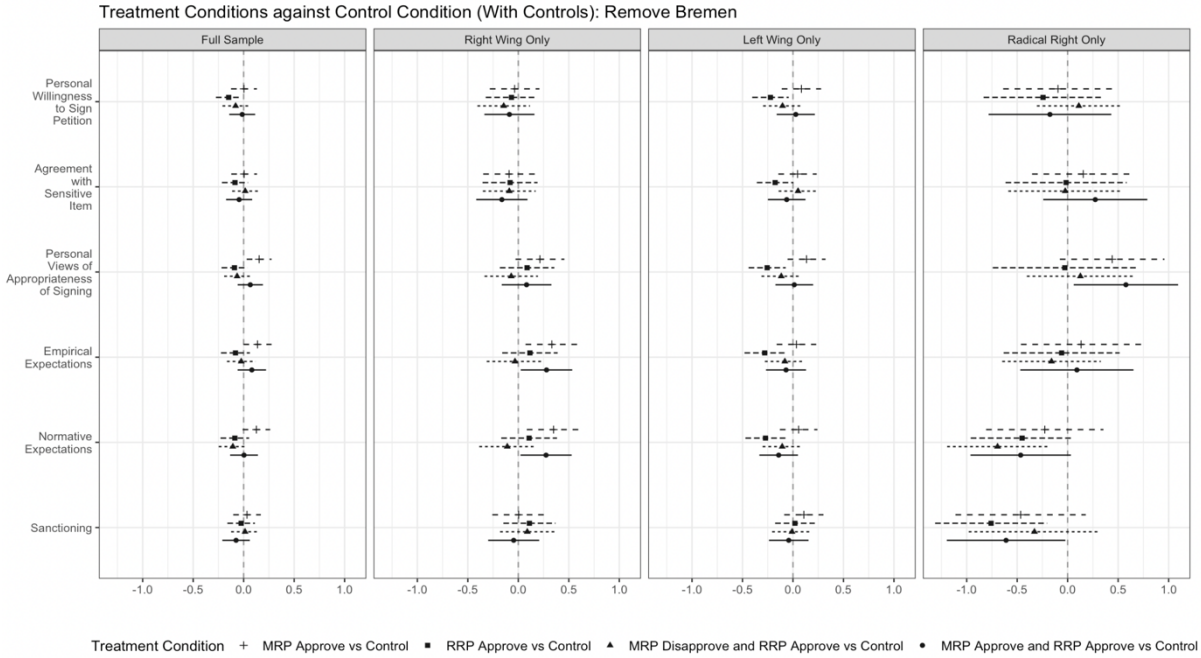


Figure F.21: Comparison of treatment and control conditions with all states except Hamburg

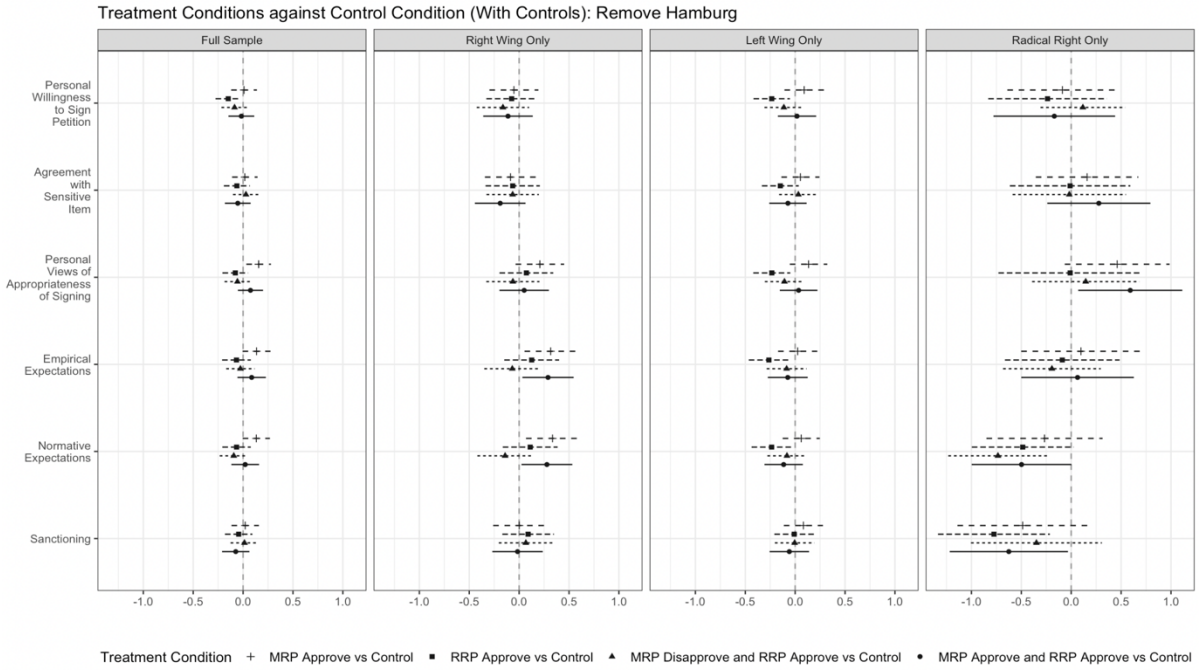


Figure F.22: Comparison of treatment and control conditions with all states except Hessen

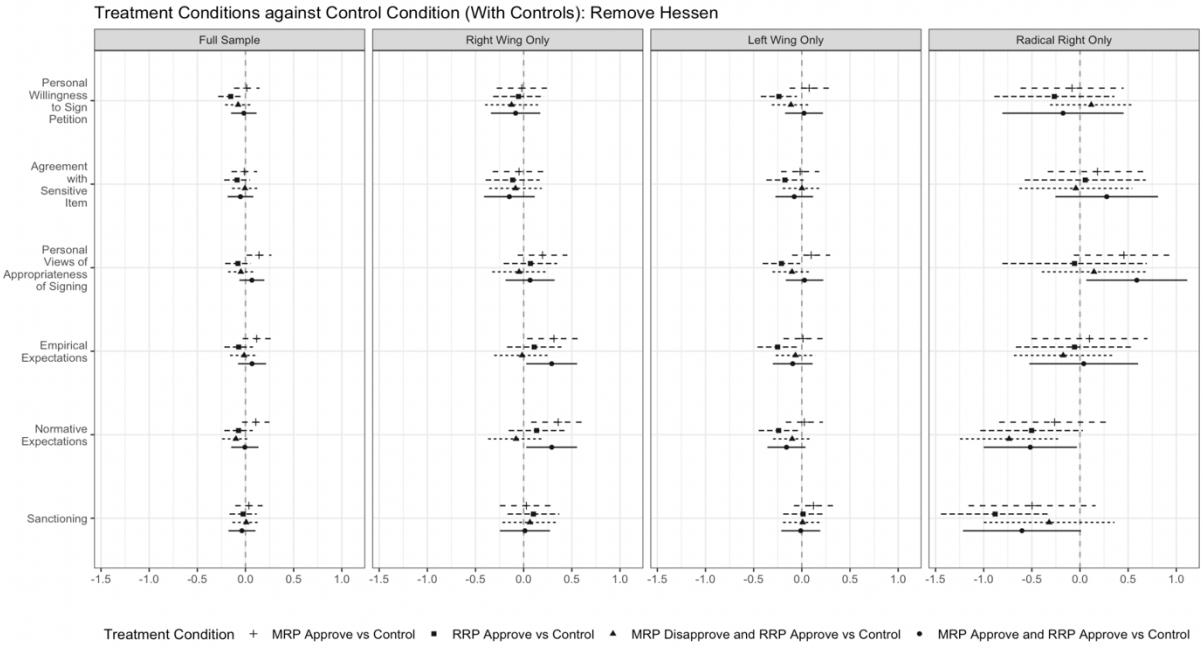


Figure F.23: Comparison of treatment and control conditions with all states except Mecklenburg-Vorpommern

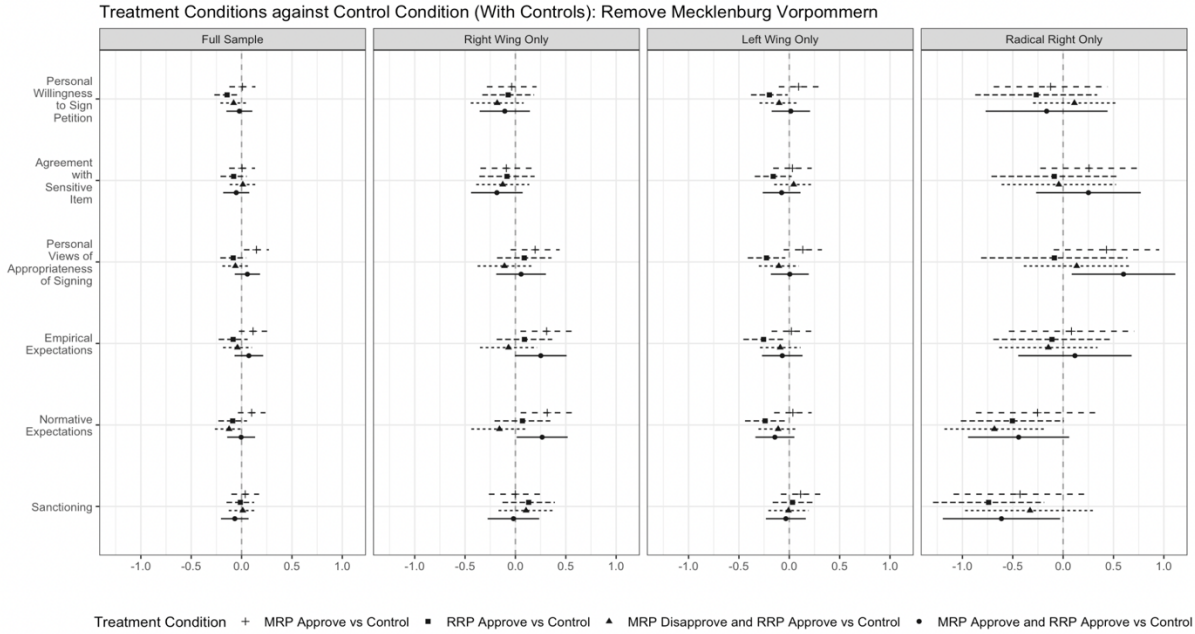


Figure F.24: Comparison of treatment and control conditions with all states except Niedersachsen

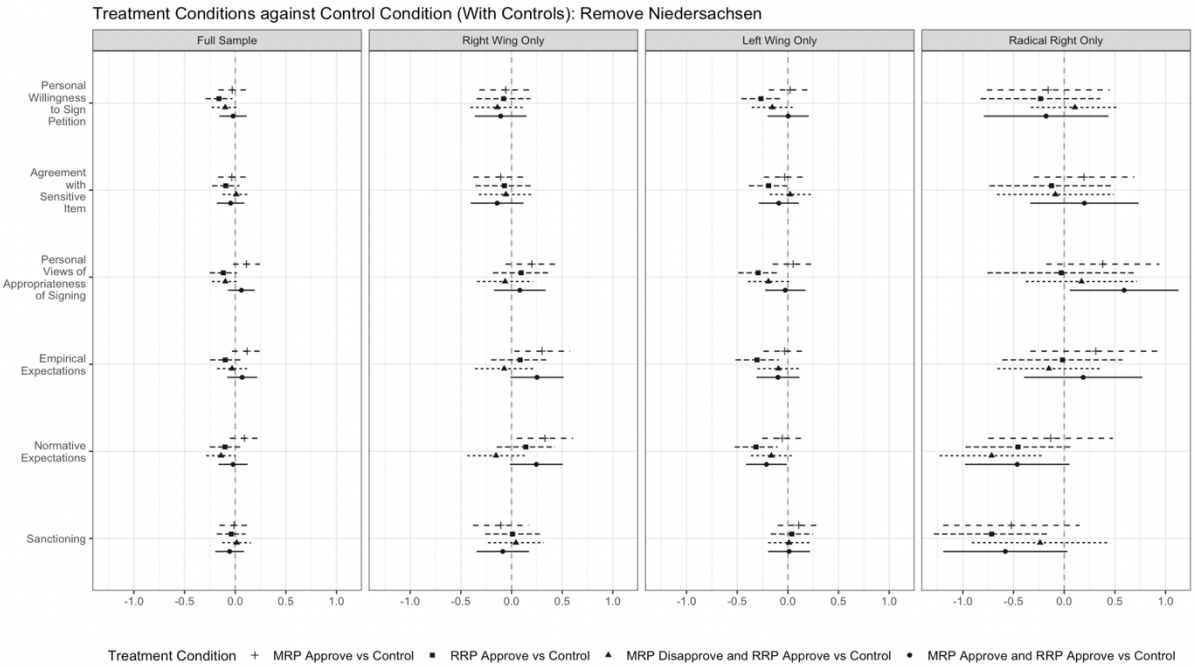


Figure F.25: Comparison of treatment and control conditions with all states except Nordrhein-Westfalen

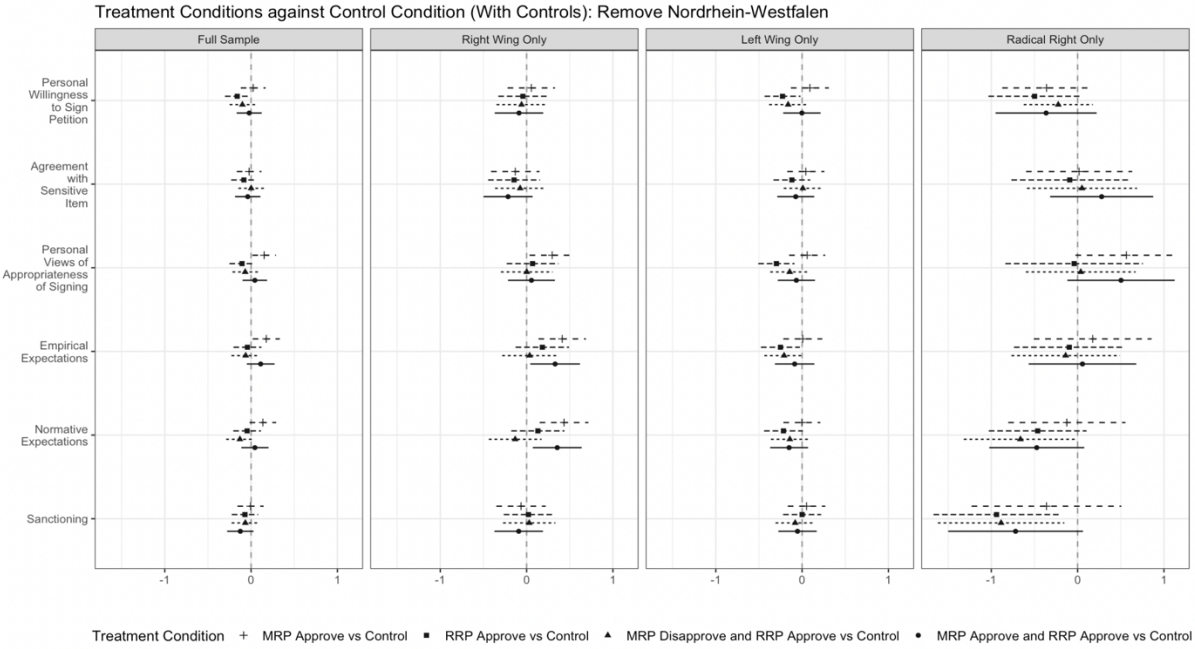


Figure F.26: Comparison of treatment and control conditions with all states except Rheinland-Pfalz

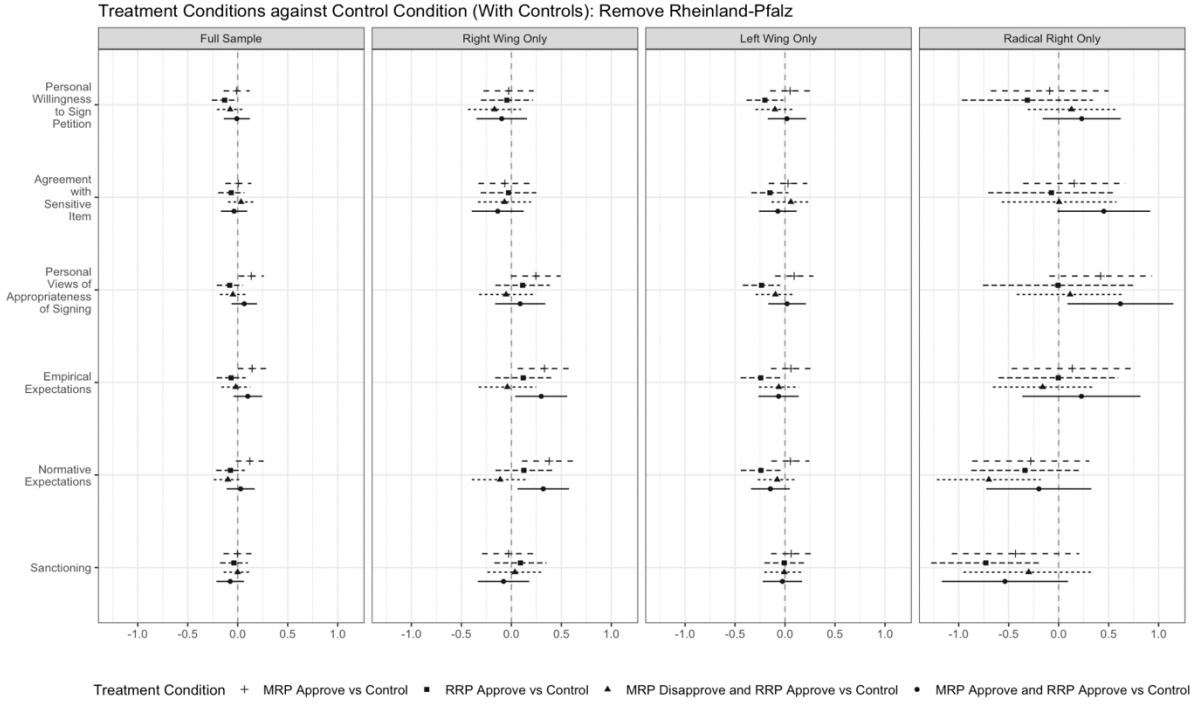


Figure F.27: Comparison of treatment and control conditions with all states except Sachsen-Anhalt

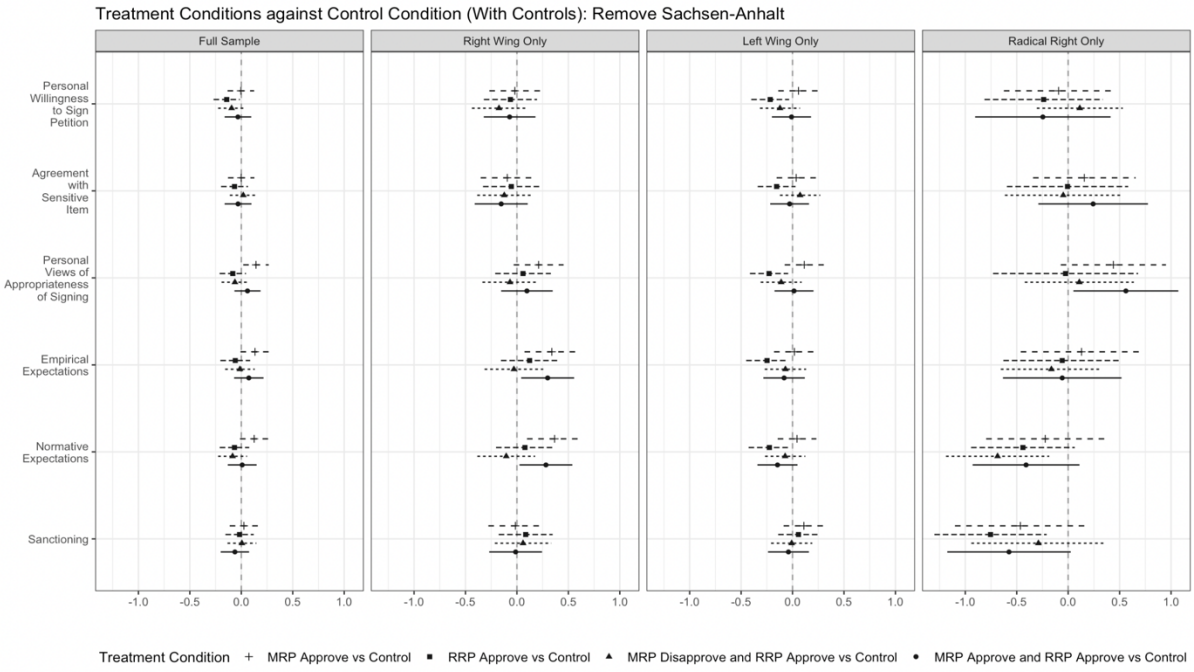


Figure F.28: Comparison of treatment and control conditions with all states except Saarland

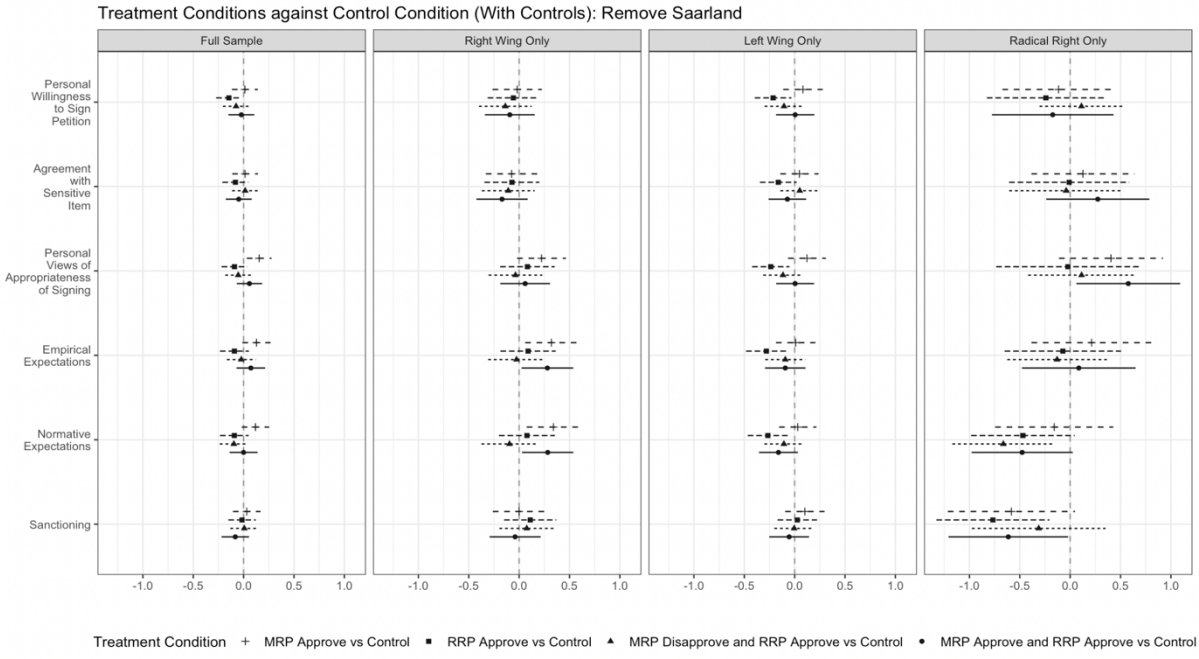


Figure F.29: Comparison of treatment and control conditions with all states except Sachsen

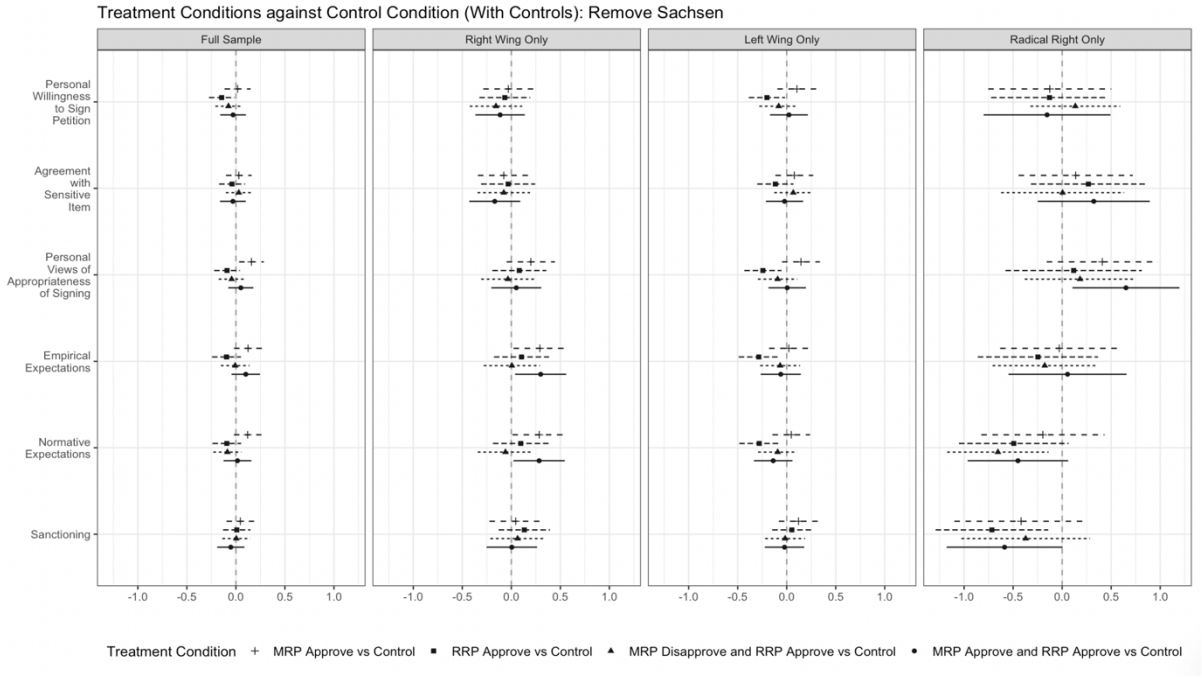


Figure F.30: Comparison of treatment and control conditions with all states except Schleswig-Holstein

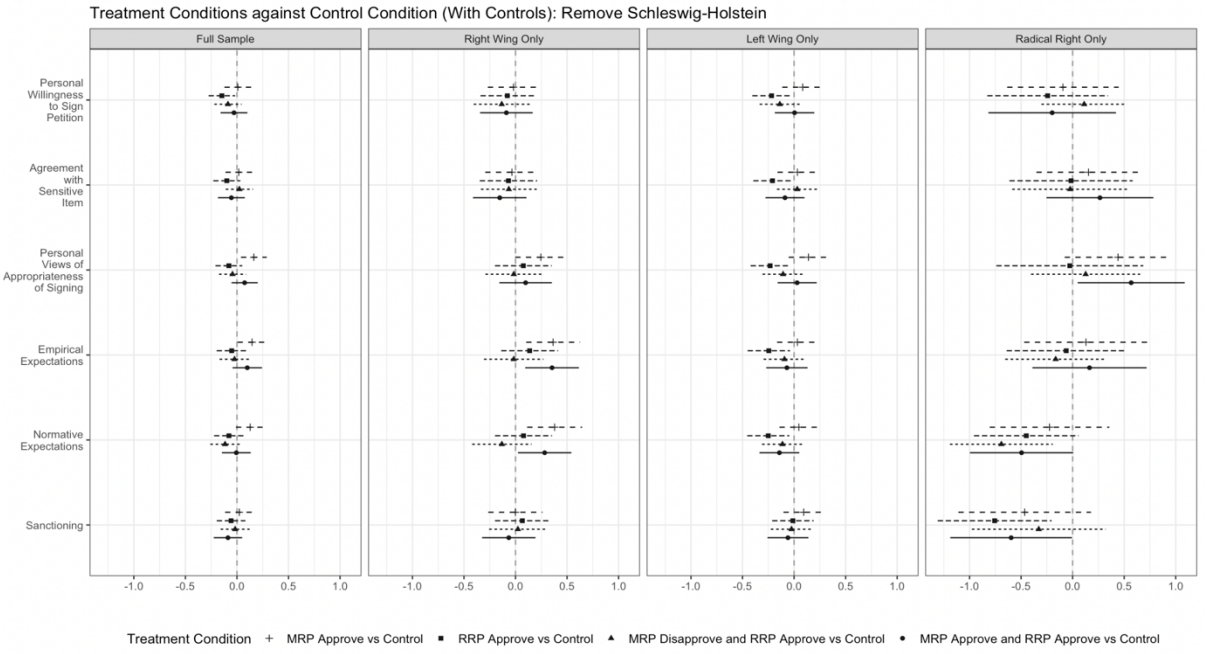
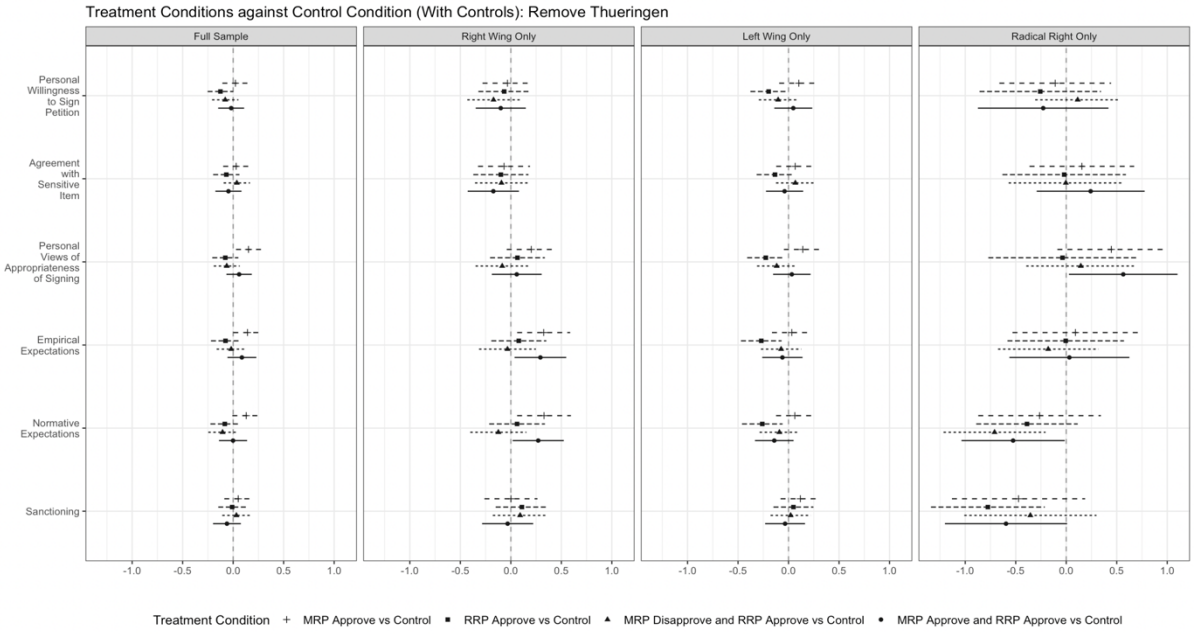


Figure F.31: Comparison of treatment and control conditions with all states except Thüringen



Appendix G: Anonymized Pre-Analysis Plan

Pre-analysis Plan for Study of the Effect of Elite Cues on the Anti-Prejudice Norm

Jan 13, 2022

Summary
The goal of this study is to assess whether elite cues from politicians alter public perceptions about the strength of the social norm against prejudice. In particular, we are interested in whether respondents exposed to xenophobic cues from mainstream politicians or radical right politicians have a weaker perception of the anti-prejudice norm, compared to those exposed to a control condition (cues on an unrelated topic).

We examine three different facets of norm perception: i) personal willingness to violate the anti-prejudice norm, ii) perception about the empirical behavior and normative beliefs of others with respect to the anti-prejudice norm, iii) willingness to sanction violators of the anti-prejudice norm.

Instrument
We field an original survey with two embedded experiments to a sample of 2000 respondents from the German population. The sample is nationally representative by age and gender, excluding individuals below the current voting age of 18, as well as those born outside of Germany. It is administered by a German survey company and fielded through an online platform. Individuals will be ascribed to one of the treatment conditions using complete random assignment (Gerber and Green 2012: 32).

The survey consists of a few components:

- Demographic information, including gender, party identification, self-placement on a political spectrum, income, state and level of education.
- Five different conditions that participants are cross-randomized into, affecting the prime they receive: *Radical Right Party (RRP) approve*, *Mainstream Right Party (MRP) approve*, *MRP+RRP approve*; *MRP disapprove + RRP approve*; and *Control.*
- A question measuring participants' willingness to express support for a xenophobic statement.
- A coordination game, where participants are shown an example of a hypothetical petition that seeks to limit immigration from the Middle East. Participants are asked about their personal willingness to sign the petition, and their perception of whether other Germans would be willing to sign the petition, and how appropriate other Germans would find signing the petition.
- A third-party punishment experiment, where participants are shown an example of a xenophobic tweet, and asked whether they would like to submit a takedown request to remove similar tweets.

Treatment
In the first two conditions (*RRP approve*, *MRP approve*), participants are exposed to vignettes which consist of quotes of anti-immigrant rhetoric, attributed to an unnamed German radical right or mainstream politician respectively.

In the third (*MRP approve + RRP approve*), participants are exposed to a similar vignette where anti-immigrant quotes are attributed to both an unnamed German radical right and mainstream politician.

In the fourth (*MRP disapprove + RRP approve*), the vignette contains quotes from a German mainstream politician that is pro-immigration in rhetoric, as well as quotes from a radical right politician that is anti-immigration in rhetoric.

To provide a control, the fifth treatment condition (*Control*) is a similar vignette on an unrelated issue, specifically, food waste reductions.

We expect our treatments to increase the personal willingness to violate the anti-prejudice norm, decrease perceptions about the empirical behavior and normative beliefs of others, and decrease willingness to sanction norm deviations. The only exception is the MRP disapprove + RRP approve treatment, which we expect will have the opposite effect.

Outcomes

Below, we detail the questions in each category that we use to construct the three outcome measures with respect to norm perception. Our empirical analyses will rely on OLS regressions of the form:

$$Y_i = \beta_0 + \beta_1 T_i + \varepsilon_i$$

where T_i is a categorical variable that indicates the treatment status of individual *i* —i.e., whether they are ascribed to the control group or to one of the treatment conditions. We will estimate this regression twice: once with no controls, and once with a full set of pre-treatment controls. We will assess the statistical significance of the estimates against a critical value of 10%. We will use heteroskedasticity-robust standard errors to derive the t-scores.

<u>Personal willingness to violate the anti-prejudice norm:</u>

Q4. Below is a statement about an issue relevant to Germany today. Would you please indicate whether you agree or disagree with it: The enrolment of pupils from Syria and Afghanistan has lowered the quality of German schools.

Q5. Many people in Germany send petitions for various everyday policy issues to the Bundestag's petitions committee. Imagine being presented with a petition that demands tighter controls and stricter enforcement of immigration from non-European countries like Afghanistan to Germany. This petition would be forwarded to the Bundestag, and the list of all signatories would be made publicly available. Would you be willing to sign this hypothetical petition?

Q6. It is appropriate to sign a petition endorsing tighter controls and stricter enforcement of immigration from non-European countries like Afghanistan.

<u>Perception about the empirical behavior and normative beliefs of others with respect to the anti-prejudice norm:</u>

Q7. Please estimate the average percentage of Germans who were willing to sign the petition when taking our survey.

Q8. Please estimate the average percentage that other Germans taking our survey indicated when they were asked how appropriate it is to sign a petition endorsing tighter controls and stricter enforcement of immigration from non-European countries like Afghanistan. The ten responses that come closest to the actual average, will receive an additional €50.

Willingness to sanction violators of the anti-prejudice norm.
Q9: Would you like to submit an anonymous takedown request for similar tweets?

Power Analyses

We have checked what the minimum detectable effect is, based on information from our pilot—as suggested in DeclareDesign blog (2019). To do so, we have run power analyses using the DeclareDesign package (Blair et al. 2019). For each outcome, we have taken the values we got in the control condition. Then, using the standard deviation that we got in each treatment condition in our pilot, we run simulations where we calculate the power for a range of potential means that we could get in the actual experiment. Table 1 below summarizes the minimum detectable effect from these analyses—i.e., the minimum effect size that we could identify with a power of 0.8. It shows that, for each treatment arm, we seem to be well powered to find many of the effect sizes yielded in our pilot.

Table 1: Minimum detectable effects in each outcome and treatment condition (power analyses conducted with Declare Design package (Blair et al. 2019))

Condition / Outcome	Personal Willingness to Sign Petition (0-1; mean in control: 0.444, σ = 0.504)	Personal Views about Appropriateness of Signing Petition (0-100, mean in control: 50.028, σ = 38.875)	Empirical Expectations about Others' Willingness to Sign Petition (0-100, mean in control: 43.333, σ = 19.621)	Normative Expectations about Appropriateness of Signing Petition (0-100, mean in control: 46.778, σ = 19.079)	Willingness to Sanction Hateful Tweet (0-1, mean in control: 0.75, σ = 0.439)
MRP and RRP approve	**0.22** 3.7 (σ = 0.506)	**8** 1.6 (σ = 35.632)	**4** 0.51 (σ = 18.613)	**4** 0.74 (σ = 17.563)	**0.22** 0.88 (σ = 0.506)
MRP approve	**0.22** 1.1 (σ = 0.485)	**7** 0.54 (σ = 32.001)	**4.5** 0.33 (σ = 22.552)	**4** 0.24 (σ = 20.785)	**0.225** 0.98 (σ = 0.505)
MRP disapprove, RRP approve	**0.22** 2.0 (σ = 0.478)	**7.5** 1.1 (σ = 32.933)	**4.5** 0.55 (σ = 22.189)	**4** 0.34 (σ = 18.670)	**0.23** 2.9 (σ = 0.478)
RRP approve	**0.22** 22 (σ = 0.504)	**7** 2.8 (σ = 32.118)	**4** 1.5 (σ = 17.954)	**4** 0.71 (σ = 19.727)	**0.22** 1.1 (σ = 0.504)

Notes: Numbers in parentheses represent the standard deviation that we got in each treatment condition in our pilot study. Numbers in bold represent the minimum effect size for which power is above the traditional threshold of 80%, assuming a sample size of *N* = 2,000 (the sample size that our experiment will have) and the standard deviation that we got in our pilot study. Numbers that are underlined represent the ratio of the estimated minimum effect size assuming a sample size of *N* = 2,000, to the effect size obtained during a pilot study with a smaller sample size of *N* = 200. Power analyses run using the multi_arm_designer function, using 500 simulations.

Additional analysis

Pairwise analyses across treatment conditions: Other than the comparison of each treatment condition to the control, to study whether there are differences in norm perception across treatment conditions, we also perform additional pairwise analyses by comparing outcomes in treatment conditions with each other.

For example, we will individually compare the outcomes in the *MRP disapprove + RRP approve* treatment condition with those in the other treatment conditions — *MRP + RRP approve*, MRP approve, RRP approve. We do the same types of pairwise comparisons with each of the other treatment conditions. These analyses will be exploratory in nature, and for that reason we do not pre-register any hypothesis regarding them.

Subgroup analysis: Using information on the respondent's gender, party identification, self-placement on a political spectrum, income, state and level of education, we will split the sample in subgroups to estimate if the effect of the treatment on norm perception is stronger among certain subgroups. These analyses will be exploratory in nature, and for that reason we do not pre-register any hypothesis regarding them.

Attrition: Should we find attrition, we will (a) test whether this is differential across groups; and (b) estimate Lee's extreme bounds.

References

Blair, Graeme, Jasper Cooper, Alexander Coppock, and Macartan Humphreys. "Declaring and Diagnosing Research Designs." American Political Science Review 113, no. 3 (2019): 838–59.

Gerber, Alan S., and Donald P. Green. *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton, 2012.

DeclareDesign blog "Should a Pilot Study Change Your Study Design Decisions?" DeclareDesign blog (2019). Accessed January 6, 2022. https://declaredesign.org/blog/2019-01-23-pilot-studies.html.

Appendix H: Consent Form For Participation in the Research Study

Study Title: Social and Political Beliefs in Germany
Principal Investigators: [*Identifying Information*]

DESCRIPTION: We are a group of non-partisan researchers interested in better understanding economic, social and political issues in Germany. This survey will take approximately 5-10 minutes. The data collected will only be used for future research.

RISKS AND BENEFITS: The risks to your participation in this online study are those associated with basic computer tasks, including boredom, fatigue, mild stress, or breach of confidentiality. In addition, you may be shown real messages from online platforms, some of which have been removed because they have been flagged for containing hateful content. The only benefit to you is the learning experience from participating in a research study and the compensation for doing so. The benefit to society is the contribution to scientific knowledge.

CONFIDENTIALITY: Please be assured that your responses will be kept confidential. The survey is administered via *[survey company name]* and the researchers will only have access to completely anonymized data. All data collected from our study will be stored securely on our servers.

COMPENSATION: You will be compensated by *[survey company name]* for taking this survey. You will not be compensated if your survey is incomplete.

SUBJECT'S RIGHTS: Your participation is voluntary. You may stop participating at any time by closing the browser window or the program to withdraw from the study. Partial data will not be analyzed.

For additional questions about this research, you may contact: [*Identfying information*]

For questions about your rights as a research participant, you may contact:

- [*Identifying Information*]

Please indicate, in the box below, that you are at least 18 years old, have read and understand this consent form, and you agree to participate in this online research study.

0. I am at least 18 years old, I have read and understand this consent form, and I agree to participate in this online research study.
    A. Yes
    B. No

If B: Thank you very much – this is the end of the survey.
If A: Move on to 1.