Supplemental Information

Re-stigmatizing the radical right: A one way street?
Balcells, Martínez, Valentim, and vanderWilden

Contents

A1 Demographic representativeness 1
A2 Balance across treatment groups 1
A3 Attrition 2
A4 Factual manipulation check 3
A5 Stigma against Vox: distributions and descriptive data 5
A6 Results consistent when including pre-treatment controls 6
A7 Measuring stigma: Index validity and individual outcomes 8
A8 Alternative estimations for conditional effects by ideology 9
A9 Treatments fail to move voting intentions for Vox 11
A10 Pre-analysis plan 12

A1 Demographic representativeness

Table A1: Comparing sample and population

	Sample Proportion	Census Proportion	Difference	P-Value
Male	0.497	0.490	0.007	0.279
18-25	0.095	0.097	0.002	0.581
26-35	0.129	0.131	0.002	0.662
36-45	0.167	0.169	0.002	0.694
46-55	0.186	0.187	0.001	0.880
56-65	0.163	0.158	0.005	0.345
66+	0.260	0.256	0.004	0.474
Andalucía	0.176	0.179	0.003	0.521
Aragón	0.029	0.028	0.001	0.792
Asturias	0.020	0.021	0.001	0.699
las Islas Baleares	0.025	0.025	0.000	0.874
las Islas Canarias	0.048	0.046	0.002	0.578
Cantabria	0.012	0.012	0.000	0.837
Castilla y León	0.050	0.050	0.000	0.917
Castilla-La Mancha	0.043	0.043	0.000	0.875
Catalunya	0.168	0.165	0.003	0.609
Valencia	0.109	0.109	0.000	0.920
Extremadura	0.022	0.022	0.000	0.958
Galicia	0.055	0.056	0.001	0.839
Madrid	0.144	0.143	0.001	0.885
Murcia	0.031	0.032	0.001	0.827
Navarra	0.013	0.014	0.001	0.694
País Vasco	0.047	0.046	0.001	0.621
La Rioja	0.007	0.007	0.000	0.912

Census proportions taken from the 2023 Spanish census (see more information here.)

A2 Balance across treatment groups

Table A2: Balance Table

	Low	High	Mean (Any Stigma)	Mean (Control)	P-value	ANOVA p-value
Age	18	80	50.37	49.93	0.40	0.61
Male	0	1	0.50	0.50	0.82	0.73
Education	1	8	5.05	4.96	0.03	0.02
Income	1	9	6.00	5.95	0.31	0.36
Political Knowledge	0	1	0.60	0.57	0.17	0.49
Spanish Identity	0	10	7.63	7.57	0.52	0.57
Regional Identity	0	10	7.86	7.68	0.04	0.04
Ideology	0	10	4.79	4.78	0.96	0.56

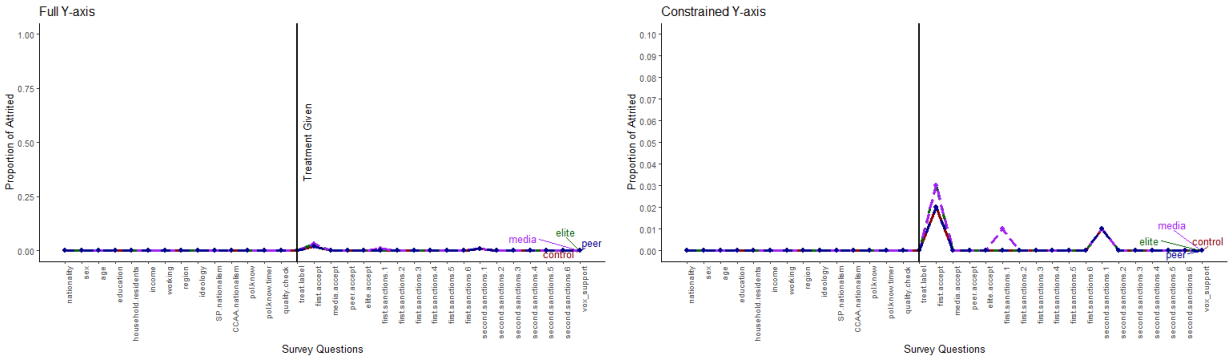
Means and p-value for combined stigma condition and control condition, ANOVA p-value for comparison of 4 groups.

A3 Attrition

We screened respondents on a factual manipulation check (pre-treatment, instructing respondents to select a specific answer). These respondents did not finish the survey (nor were fully compensated) and were returned to the survey firm's site (n = 2817). We also removed "speeders" (less than 90 seconds to complete the full survey, n = 34) and "slow" respondents (greater than 30 minutes to complete the full survey, n = 53) from the analysis. On average, these removed respondents were younger, less educated, lower income, and more likely to be female, compared to the final sample.

We also excluded those who did not finish the survey from our analysis (n = 203, a relatively small number compared to the full sample of 5526). Figure A1 shows where people dropped out in the survey (the right panel constrains the y-axis for greater readability). The first thing to note is that, overall, attrition is quite minor and does not appear to be problematic. Second, while we do observe a very small level of post-treatment attrition, we notice that attrition does not significantly differ between treatment assignment, thus reducing the threat of attrition bias. In general, we are not particularly worried that our results are begin unduly shaped by attrition or attrition bias.

Figure A1: Survey Attrition



A4 Factual manipulation check

We included a question asking: "Which of the following best describes **how the previous text** discussed Vox?" (1) Stigmatized; (2) Neither stigmatized nor normalized; (3) Normalized. Among the treated, only 69% or respondents correctly identified the text as "stigmatizing," despite all of these respondents correctly passing the instructional manipulation check prior to the treatment. Figures A2, A3, and A4 replicate Figures 2, 3, and 4 from the main text when excluding respondents who did not identify the text as "stigmatizing" from the analysis.

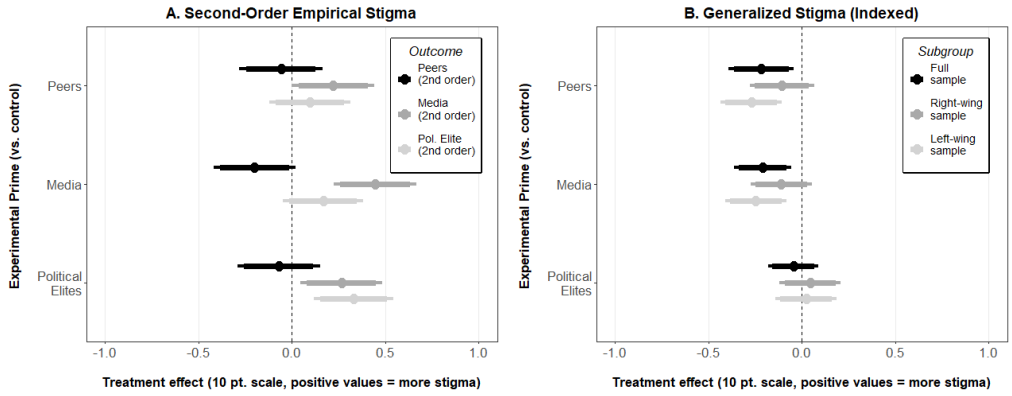


Figure A2: Replicating Figure 2, excluding incorrect identification respondents

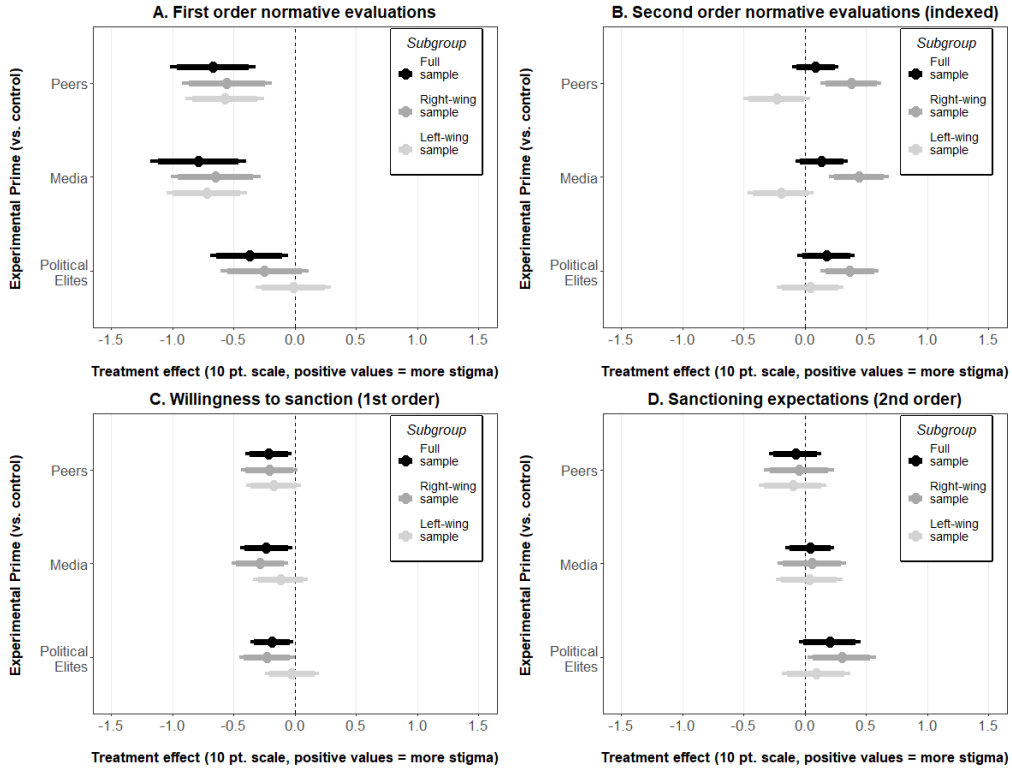


Figure A3: Replicating Figure 3, excluding incorrect identification respondents

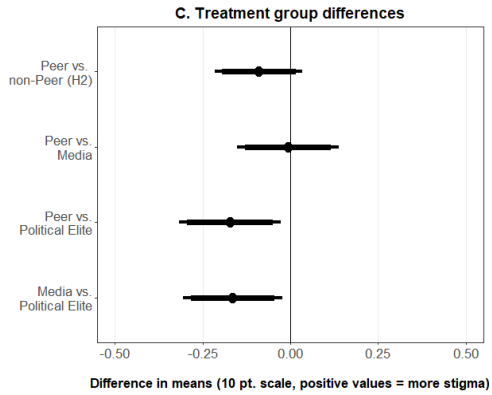


Figure A4: Replicating Figure 4, excluding incorrect identification respondents

When we exclude respondents who did not identify the text as "stigmatizing" from the analysis, our results remain largely consistent, with some exceptions. In generally, these results suggest more consistency than divergence, with the overall conclusions of null and possible backlash effects holding.

- Figure A2A shows that the substantive manipulation check was successful for the "media" and "political elite" primes: when told about how these sources stigmatized Vox, individuals report lower acceptability of Vox among that group (2nd order acceptability). We do not see the same movement for the "peer" cue, though (which reflects the same conclusion presented in the main text). Figure A2B mirror those in the main text (Figure 2B).

- Figure A3 shows that largely similar results to Figure 3. However, we do identify larger backlash effects on first-order normative evaluations, as well as some (inconsistent) evidence of the same backlash reflected in first-order willingness to sanction.

- Finally, we note some small differences in results when comparing Figure A4C to Figure 4. Among the full sample, we find support for H2. However, the results appear slightly more inconclusive when examining the "surely treated" sample. This is likely because the "media" prime appears to elicit a similar (compared to the "peer" prime) backlash.

A5 Stigma against Vox: distributions and descriptive data

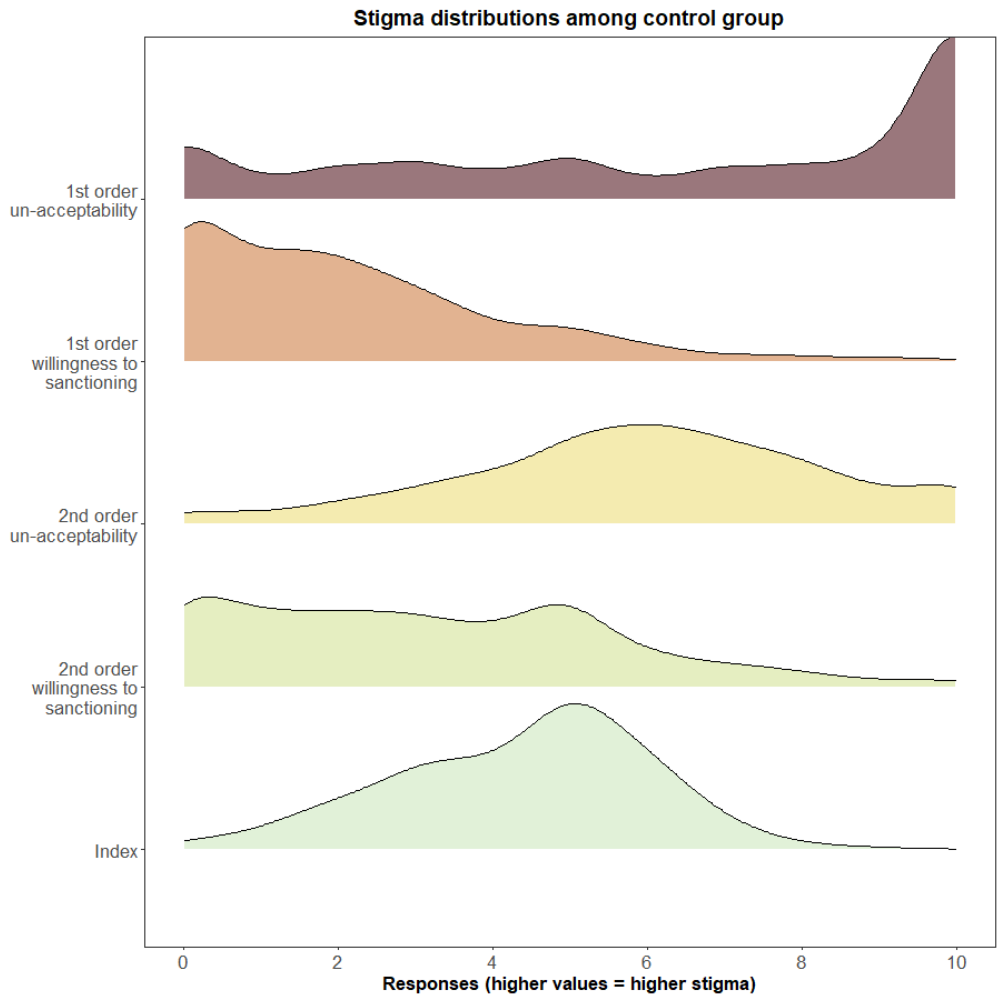


Figure A5: Descriptive data: Stigma towards Vox

Table A3: Descriptive statistics

	Mean (control)	SD (control)	Mean (full)	SD (full)
Generalized Index	4.39	1.65	4.31	1.61
1st order (un)acceptability	6.36	3.68	6.1	3.64
1st order sanctioning	2.08	2.02	2.06	2.02
2nd order (un)acceptability	5.91	2.35	5.93	2.22
2nd order sanctioning	3.18	2.41	3.16	2.38

A6 Results consistent when including pre-treatment controls

Models in Figures A6, A7, and A8 largely mirror those used for Figures 2, 3, and 4, when controlling for age, gender, education, income, political knowledge, Spanish identity, territorial identity, and including regional fixed effects.

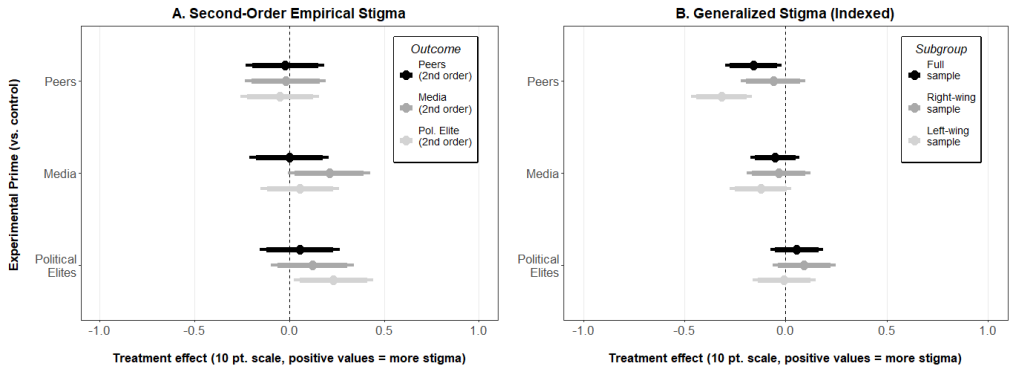


Figure A6: Replicating Figure 2 when including controls

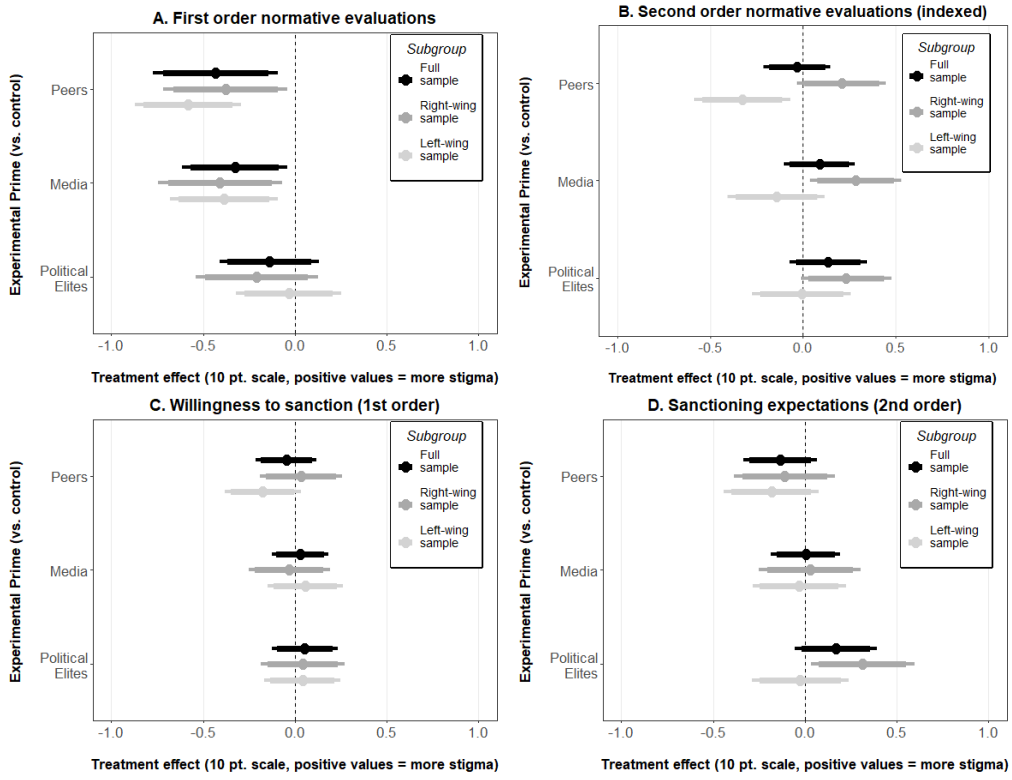


Figure A7: Replicating Figure 3 when including controls

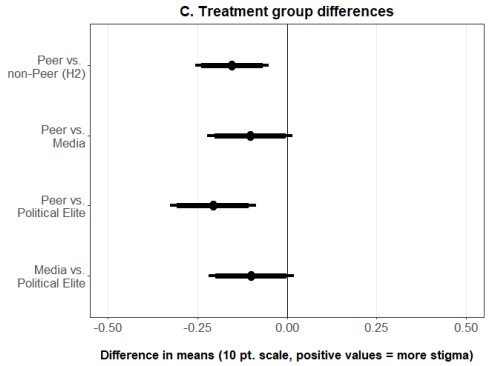


Figure A8: Replicating Figure 4 when including controls

Similar to the analysis in Section A4, these alternative model specifications suggest an overall picture of consistency, with few exceptions noted:

- First, when we include controls for models in Figure 2 (Figure A6), we notice that the substantive manipulation check for the "political elite" prime appears to work is statistically significant at the 0.05 level).
- Second, when we include controls for models in Figure 3 (Figure A7), the general picture looks quite similar – most of the movement is in normative evaluations. For the "peer" and "media" primes, we see a relatively uniform backlash, where individuals in the treated group view Vox as more acceptable. Like before, we find heterogeneous effects on second-order normative evaluations, with those on the right largely accepting the manipulation as intended, while those on the left appear to reject the treatment. When including controls, we only find statistically significant effects on the left for the "peer" treatment group, though. Finally, we again see little movement on willingness to sanction (with one exception of right-leaning respondents perceiving of second-order sanctioning as more likely following the "political elite" prime.
- Third, when including controls for models from Figure 4 (Figure A8), we notice that the comparison between the Media stigma and Political Elite (row 4) becomes statistically significant at the 0.10 level. The rest of coefficients and p-values remain virtually equal to Figure 4.

A7 Measuring stigma: Index validity and individual outcomes

Should we measure generalized stigma as a bundle of 1st and 2nd-order normative evaluations and sanctioning expectations? We argue that these each capture relevant angles of generalized stigma. In this section, we empirically show how and which elements of our index generally map on to one another. Table A4, which relies on data exclusively from the control (or, untreated) group suggests that, despite our prior expectation that these variables would generally move together, inter-correlations between each item remains relatively low. However, we identify positive and relatively high correlations between the normative evaluations (1st and 2nd order) and between the sanctioning expectations (1st and 2nd order).

Table A4: Inter-correlations in generalized stigma index

	1st (un)-acceptability	2nd (un)-acceptability	1st sanctioning	2nd sanctioning
1st (un)-acceptability	1			
2nd (un)-acceptability	0.49	1		
1st sanctioning	0.23	-0.09	1	
2nd sanctioning	-0.01	-0.07	0.49	1

When using a principal component analysis to examine responses to each indexed item, we similarly identify a difference between the normative items and the sanctioning items. The first two dimensions from the PCA, which account for about 73% of the overall variance, and their corresponding factor loadings, are shown in Figure A9.

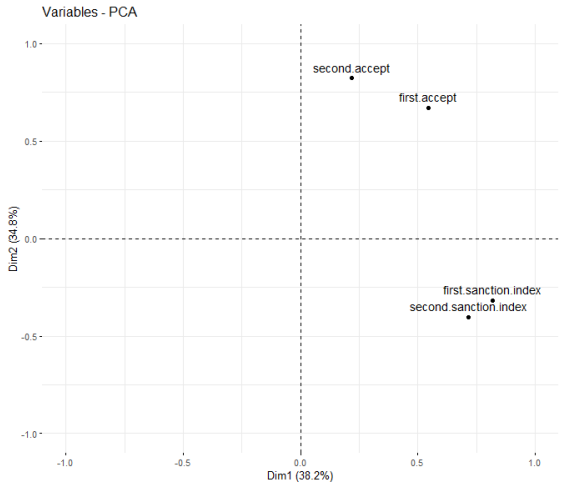


Figure A9: Principal component analysis for stigma index (first 2 dim)

The figure suggests that, while all factors load onto the first dimension in the same direction, we see opposite factor loadings on the second dimension when looking at the normative versus sanctioning factors. Given these analyses, and following our stated roadmap from our pre-registration, we include an analysis of individualized outcomes in the main paper.

A8 Alternative estimations for conditional effects by ideology

In our main figures, as well as multiple appendix figures, we separate the sample into "right" and "left" wing respondents. We include a pre-treatment question in our survey asking respondents to place themselves on an ideological scale between 0 and 10, where 0 indicates that you are far on the left and 10 indicates that you are far on the right. To present our results, we roughly split the sample, where "left" wing respondents are those identifying between 0:4, and "right" wing respondents are those identifying between 5:10, as specified in our pre-analysis plan. Under this cutoff, we identify a statistically significant difference treatment effects for "left" and "right" wing respondents for the peer-cued stigma group, when compared to the control group, and the interaction term approaches statistical significance ($p < 0.10$) for the Media-cued group. Table A5 shows the results of these interaction models.

Table A5: Interaction: Treatment effects on generalized stigma by ideology

	Peer-cued	Media-cued	Elite-cued
(Intercept)	5.276*	5.276*	5.276*
	(0.058)	(0.057)	(0.057)
treated	-0.292*	-0.198*	-0.046
	(0.082)	(0.081)	(0.081)
right	-1.626*	-1.626*	-1.626*
	(0.078)	(0.077)	(0.077)
treated × right	0.225*	0.186†	0.121
	(0.111)	(0.109)	(0.109)
Num.Obs.	2768	2761	2745
R^2	0.217	0.224	0.232

† $p < 0.10$, * $p < 0.05$

We recognize that this is a somewhat arbitrary cutoff. To assess the sensitivity of our choices, we also ran two additional specifications: using a 3-level cutoff and treating ideology as continuous. Table A6 shows that the significance of the interactions is sensitive to modeling choices. As such, we temper our language in the main text, framing our findings in suggestive rather than definitive terms.

Table A6: Alternative ideology interaction models

	Peer-cued		Media-cued		Elite-cued	
	3 levels	Continuous	3 levels	Continuous	3 levels	Continuous
(Intercept)	4.107*	5.858*	4.107*	5.858*	4.107*	5.858*
	(0.074)	(0.074)	(0.073)	(0.073)	(0.072)	(0.073)
treated	-0.120	-0.264*	-0.071	-0.176†	0.057	-0.023
	(0.104)	(0.107)	(0.101)	(0.105)	(0.102)	(0.104)
left	1.169*		1.169*		1.169*	
	(0.093)		(0.092)		(0.091)	
right	-0.878*		-0.878*		-0.878*	
	(0.103)		(0.101)		(0.100)	
treated x left	-0.172		-0.128		-0.103	
	(0.131)		(0.128)		(0.129)	
treated x right	0.094		0.022		0.033	
	(0.145)		(0.143)		(0.141)	
ideology		-0.308*		-0.308*		-0.308*
		(0.013)		(0.013)		(0.013)
treated x ideology		0.019		0.013		0.007
		(0.019)		(0.019)		(0.019)
Num.Obs.	2768	2768	2761	2761	2745	2745
R^2	0.253	0.259	0.263	0.265	0.272	0.281

† $p < 0.10$, * $p < 0.05$

A9 Treatments fail to move voting intentions for Vox

Footnote 2 notes that, as we intend to alter perceived norms around Vox, it is also possible that we could effect voting behavior surrounding the party. We use a five point scale of "likelihood of future voting for Vox," where respondents were asked whether they would ever vote for Vox in the future. Response options included: (1) I will never vote for Vox; (2) it is unlikely that I would ever vote for Vox; (3) I'm not sure if I will ever vote for Vox; (4) I will likely vote for Vox in the future; (5) I am sure that I will vote for Vox in the future. Figure A10 shows that the primes fail to move respondents, including respondents on the "political right," on future intentions to vote for Vox.

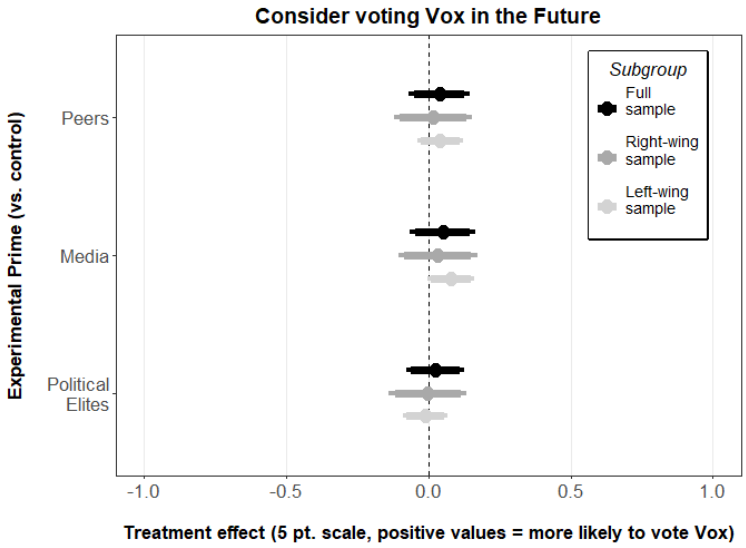


Figure A10: Primes fail to move future voting intention for Vox

A10 Pre-analysis plan

Stigmatizing the radical right

June 14, 2024

Abstract

Radical right behavior and support for radical right parties have increased across many countries in recent decades. A growing body of research has argued that, similar to the spread of other extremist behaviors, this is due to an erosion of political norms against supporting these parties. This would suggest that strengthening norms by stigmatizing radical-right parties might be an effective way of countering their growth. But what interventions are most likely to increase such stigma? To shed light on this matter, we propose a survey experiment in Spain that compares the effectiveness of a number of theory-driven interventions aimed at increasing political stigma against a radical right party. Our findings will shed light on the efficacy of particular primes of stigma, informing theoretical and policy debates on the determinants of radical right behavior.

Word count: 3928

1 Introduction

An emerging body of literature argues that the increasingly frequent strong electoral performances by radical right parties (Mudde 2016), as well as the rise of counter-normative extreme behaviors like hate crimes (FBI 2023), are partially triggered by changing social norms. According to this argument, the normalization of radical right parties has paved the way for increased expression of counter-normative preferences, behaviors, and endorsements of extremist political parties (Tankard and Paluck 2016; Bursztyn, Egorov, and Fiorin 2020; Valentim 2021; Alvarez-Benjumea and Valentim 2024; Gul 2023; Albornoz, Bradley, and Sonderegger 2022; Romarri 2020). Strengthening anti-radical right norms, then, could be an effective way of reducing radical-right behavior (Ammassari 2024).

What interventions strengthen anti-radical right norms? Norms depend on one's perceptions of what is deemed acceptable in their social environment (Bicchieri 2016). Since these second-order expectations are typically not directly observable, citizens rely on cues that help them generate norm perceptions. We argue that identifying and manipulating those cues can be a way of strengthening norms against radical-right behavior.

Existing literature has highlighted three sources of cues that might affect norm perceptions. First, cues can come from the behavior of other citizens that one interacts with in their daily life (Blinder, Ford, and Ivarsflaten 2013; Harteveld et al. 2019; Oshri et al. 2023; Alvarez-Benjumea and Valentim 2024; Ammassari 2022). Second, cues can emanate from the media (van Heerden and van der Brug 2017; van Spanje and Weber 2019; Bolet and Foos 2023). Finally, cues may be signaled by the behavior of political elites (Clayton et al. 2021; de Jonge and Gaufman 2022; Krause, Cohen, and Abou-Chadi 2023; Axelsen 2023). Experimentally, researchers have used interventions from these actors to increase (or decrease) the salience of social stigma surrounding radical right parties (for example, Harteveld et al. 2019; Bursztyn, Egorov, and Fiorin 2020), though there is little systematic validation of these primes.

Following this literature, we build and validate experimental treatments designed to manipulate stigma and assess their effectiveness in strengthening anti-radical right norms. Our proposed experiment will randomly assign a representative sample of 5,500 Spanish participants to four groups corresponding to the three stigma-inducing sources identified in the literature (peers, media, and political elites) and a control group that receives a neutral prime about the radical-right party Vox. Following the experimental treatment in which a source cues stigma against Vox, participants will answer a series of questions assessing the strength of the political stigma surrounding the party, including items that capture first and second order perceptions of the acceptability of Vox and the likelihood of sanctioning public displays of support for Vox (Bicchieri 2016).

Our project has two main goals: (1) to validate source-specific primes of political stigma, assessing whether and how a given prime moves anti-radical right norms, and (2) to assess which cues are more effective in strengthening anti-radical right norms. Written for clarity, parsimony, and generalizability, our primes are easily adaptable to different cases. As such, our project will be useful to experimentalists working on radical-right parties and will speak to political debates on how to best strengthen norms against radical-right behavior.

2 Theoretical focus and hypotheses

We aim to test whether source-specific primes stigmatizing the radical right move generalized perceptions of stigma by examining the effect of each treatment on a summary index of stigma, which includes first and second order evaluations of radical-right parties' acceptability and the likelihood of sanctioning for publicly supporting such a party.[1] For each

[1]We will use principal component analysis to examine the degree to which our outcomes move together. If, as we expect, our outcomes co-vary as a single factor, we will use that factor as an outcome variable. Otherwise, we will present results of the separate factors. Additionally, we will present each outcome separately in the appendix (see Figure A1).

prime of stigma, we will test the hypothesis:

- H1A: Priming a source-specific stigma will increase generalized stigma of Vox.

It is also possible that the effects of our treatments on generalized stigma, which include an individual's personal normative dispositions, will be conditional on ideology. Past evidence documents that leftists often already deem radical-right parties unacceptable (Golder 2016; Dinas, Martínez, and Valentim 2024). If this holds, we may encounter a "ceiling effect," as it could be difficult to shift individual perceptions of the appropriateness of the radical right, while centrist and right-wing individuals may be more "movable." Therefore, we will explore the hypothesis:

- H1B: Among respondents ideologically placed in the center or right, priming a source-specific stigma will increase generalized stigma of Vox.

Nonetheless, we treat H1B as an exploratory hypothesis, as our main quantity of interest is an average treatment effect, rather than a treatment effect conditional on ideology. This is because the entire population, which includes individuals across the political spectrum, plays a role in constituting and upholding norms. Therefore, strengthening perceived stigma against the radical-right among left-wing, centrist, and right-wing individuals can all affect the expected prominence of radical right behavior in the future.[2] Furthermore, by using a sampling strategy that does not target specific ideological types, we are able to see to what extent these different movements—if they exist at all—cancel each other out.

While the main task of this project is to validate experimental manipulations of stigmatization efforts against the radical right, we are also interested in comparing the efficacy of different source cues. Recent observational evidence documents mixed results concerning

[2]We would expect that treatments are most likely to affect voting behavior among right-wing individuals, who are least likely to have ruled out voting for Vox completely. We include a survey question asking about an individuals' likelihood to vote for Vox which can be used for additional exploratory analysis.

the effects of elite-led acceptance or stigmatization strategies, such as coalitions or *cordon sanitaires* (van Spanje and Weber 2019; Favero and Zulianello 2023). Furthermore, peer-based stigmatization may be more efficient in deterring radical right support, as peers are the more proximate and relevant reference group in most individuals' everyday lives (Ammassari 2024). Accordingly, cuing stigma via a vignette related to one's peers may have stronger effects than media-led or political elite-led stigmatization.

- H2: Stigma cued via peers is more effective in increasing generalized stigma than stigma cued via media or political elite sources.

Validating and comparing the efficacy of treatments is our ultimate goal, rather than testing a specific theory. Accordingly, we will additionally test differences in means between each possible dyad, though we power off of successfully identifying H1A and H2 (4 total tests).

3 Research design

To validate and compare the efficacy of various sources of stigma, we plan to run a four-arm survey experiment in Spain.[3] Our treatments will manipulate political stigma concerning Vox, the first radical-right party to obtain parliamentary representation after the Francoist dictatorship. Despite being the third-largest party in the Spanish Congress, Vox is widely perceived as a party that challenges the fundamental principles of liberal democracy and is deemed unacceptable by a significant segment of the electorate.[4] This status of partial normalization and stigmatization offers necessary latitude to reasonably manipulate the perceived normative expectations associated with Vox using a survey instrument, as

[3]The sample will include 5,500 online respondents recruited with the firm 40dB.

[4]This description aligns with qualitative and quantitative evidence on the Swedish Democrats (SD), the Italian League, or the Swiss People's Party (SVP) (Favero and Zulianello 2023; Ammassari 2024).

argued by recent research (Alvarez-Benjumea and Valentim 2024).

Figure 1: Consort diagram

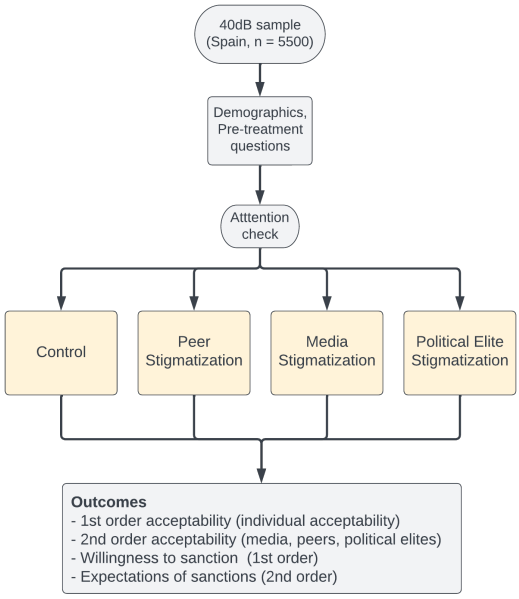


Figure 1 depicts the structure of our experimental design. The instrument firsts gathers socio-demographic and ideological profiles of respondents before random assignment into one of four treatment arms.[5] In the control condition, respondents receive a neutral message that primes them to think of Vox, void of information related to social norms surrounding the party. We include three treatment groups that each represent a specific source cuing social stigma against Vox. Treatment arms are structured to be as similar as possible to ensure information equivalence and vary only the source of the stigma. Table 1 displays the text included in each treatment arm.

[5]See Appendix A1 for a draft of the survey instrument.

Table 1: Vignettes corresponding to each treatment arm.

Control	Peer Stigmatization	Media Stigmatization	Political Elite Stigmatization
Vox is a political party that was formed in 2013.	Vox is a party that is considered to be **very stigmatized** by *citizens in this country*. For example, if someone were to publicly show support for Vox or one of its proposals, whether on the street or in a public square, it is highly likely that other peers or people passing by, both right- and left-wing, would judge or marginalize that individual.	Vox is a party that is considered to be **very stigmatized** by *the media in this country*. For example, if a journalist or political commentator were to publicly support Vox or one of its proposals, it is highly likely that other journalists or commentators, both right- and left-wing, would judge or marginalize that commentator.	Vox is a party that is considered to be **very stigmatized** by *politicians in this country*. For example, if a politician were to publicly express support for Vox or any of its proposals, it is highly likely that other politicians, both right- and left-wing, would judge or marginalize that politician.

After respondents are exposed to the treatment, they will respond to questions tapping first and second order perceptions of political stigma surrounding the party. The presentation of blocks and questions within blocks will be randomized.[6] One possible concern is that responses to our survey items are affected by preference falsification (Kuran 1995; Valentim 2024). In making respondents more aware of norms, our treatments could make them insincerely respond to outcome items. However, preference falsification in survey item responses can be seen, itself, as a measure of the social norms in place and their strength (Valentim 2021). As such, even if the responses provided are to some extent affected by this phenomenon, that would still suggest that our treatments are increasing political stigma.

Our survey will include the following outcomes:

- **First order evaluations**
  - On a scale of 0:10, how acceptable do you think it is to support Vox?
- **Second order evaluations**

[6]Section A1 of the Appendix includes the Spanish translation of the different survey items.

  - On a scale of 0:10, how acceptable do you think Vox is among most Spanish citizens?
  - On a scale of 0:10, how acceptable do you think Vox is among most of the Spanish media?
  - On a scale of 0:10, how acceptable do you think Vox is among most Spanish political elites?
- **Willingness to sanction (First and Second Order)**
  - Imagine that someone in a public square is wearing a T-shirt with Vox's logo. How likely (on a scale of 0:10) would you be to have one of the following reactions/Now imagine someone else is in the same situation.How likely (on a scale of 0:10) would you expect that person to have one of the following reactions]:
    * Call out the individual
    * Insult the individual
    * Act physically aggressive towards the individual
    * Gossip about that individual
    * Deny help to that individual
    * Avoid interactions with that individual

3.1 Testing hypotheses

To test our hypotheses, we plan to estimate the average treatment effect of cuing source-specific stigma on generalized stigma by using a simple difference-in-means estimator.[7] As a substantive manipulation check, we will examine the differences-in-means between a source-specific stigma prime and the control condition on the corresponding second-order expectation. For example, we will examine the effect of media stigmatization on second order expectations of media-induced stigma against Vox.[8]

[7]We also plan to assess balance on observable covariates, and can use covariate-adjusted OLS modeling as a robustness check.

[8]The substantive manipulation is also part of our outcome index, as perceptions of second-order normative evaluations are relevant to the general stigma surrounding a party. We also include a factual manipulation check that simply asks if the previous text described Vox as stigmatized, normalized, or neither. Including this allows us to be sure that respondents understood that the (treatment) text describes Vox as stigmatized.

To test H1A and H1B, we will calculate the difference-in-means between a source-specific stigma prime and the control group on an indexed outcome of generalized stigma, operationalized as the average stigma across the four outcomes.[9] We will test H1B by conditioning the sample by those identifying between a 5:10 on the self-reported ideological scale (measured pre-treatment).[10]

To examine which treatments increase stigma the most (H2), we will compare the difference in generalized stigma among the peer-cued treatment group and a combined media and political elite-cued treatment group. Our intuition is that peers are a more relevant reference group than the media or political elites, meaning that the peer-based prime is more effective than the remaining treatments. As exploratory analysis, we also plan to report the differences between pairwise comparisons of these primes (see Figure A1 for an example).

In total, we power off of detecting significant effects from 4 main hypotheses (3 tests for H1A, 1 test for H2). As we expect effect sizes for H1B to be larger than ATEs identified in H1A, we select not to power off of H1B, and treat this as an exploratory analysis into the heterogeneity and scope of stigmatization attempts.[11] We plan to reduce the threat of type II errors by using Benjamini-Hochberg corrections for the 4 pre-registered hypotheses. Additionally, we will address missingness and attrition by empirically diagnosing and discussing threats posed by these challenges, and possibly using weighting techniques to correct for issues, should they arise (which we do not expect).

[9] As discussed in Footnote 1, we will also analyze whether it is appropriate to treat these outcomes as a single factor, and present disaggregated results in our appendix.

[10] We also plan to explore the sensitivity this choice, using operationalizations like 3-level specifications of leftists, centrists, and rightists or treating ideology as a continuous variable to be interacted with the treatment.

[11] Additionally, we leave open the possibility of further exploratory analyses, which will be reported in the appendix. These include probing possible heterogeneous effects by age, gender, education, nationalism, and geography.

3.2 Simulated data

To show the plausibility of our design and the intended structure of our analysis, we simulate data for a sample of 5,500 respondents. We simulate the data on a (conservatively) hypothesized effect size of at least 0.2 (Cohen's d) for each pre-registered hypothesis.[12] Let us assume that H1A and H2 are fully supported, with an effect size of at least 0.2 for priming a source-specific stigma on generalized stigma, and an effect size of 0.1 for the difference between peer-cued stigma and media/political elite-cued stigma.

Simulating these data in `R`, we plot 90 and 95% confidence intervals for the effect of each prime compared to the control group (as we will do with the final data). The left panel shows our substantive manipulation check: that source-specific primes of stigma move second order expectations of stigma for the corresponding group. The right panel shows the main tests for H1A and H1B.[13]

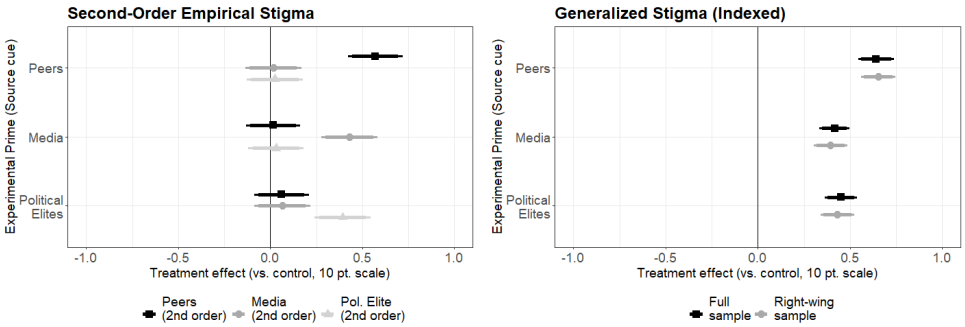
Given these data, we would claim support for H1A (and H1B). Furthermore, the simulated findings shown in the left panel would indicate that our base-level substantive manipulation check (on second-order evaluations) was consistent with our expectations. Finally, these data would also support H2, suggesting that peer-cued stigmatization is more effective than media/political elite-cued stigmatization (difference = 0.1988, SE = 0.0326, p-value $<$ 0.001).[14]

[12]This effect size mirrors that found in a loosely related manipulation check done by the authors in a previously run survey (currently under review). Moreover, given the directness of our experimental primes and outcome measures, we anticipate that effect sizes may be larger when running our tests.

[13]Estimates for the full sample include confidence intervals corrected with critical t-values for significance using the Benjamini-Hochberg method across four hypotheses. Estimates on first order evaluations and willingness to sanction are presented in Figure A1.

[14]Figure A1 plots the pairwise comparisons of each treatment and treatment effects on the

Figure 2: Simulated results



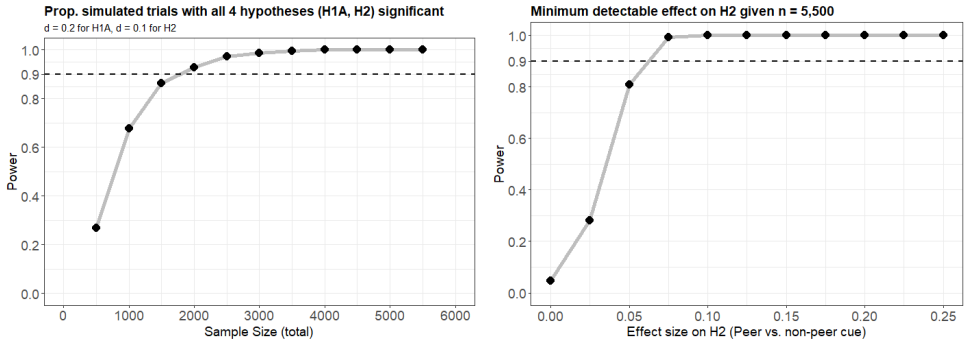
Such results would indicate two major takeaways. First, simple primes that cue stigma against the radical-right would be validated. Second, they would suggest heterogeneity in the ability of different cues sources to increase stigma against the radical-right. Specifically, these results show that peer-cues induces greater stigma than media- or political elite-cues.

Our sample size is based off of available resources and existing contracts with a survey firm. It is also empirically justified. If we assume that an effect size of 0.2 is a reasonable (and likely, conservative) effect to power off of, we are well powered to detect significant effects on the four comparisons (H1A and H2) discussed above. The left panel of Figure 3 plots the proportion of simulations under the above assumptions that produce statistically significant results on all four hypotheses (with a critical threshold of 0.05, corrected for four hypothesis tests using the Benjamini-Hochberg method) out of 1000 simulations at each n-size. As the figure suggests, we would be well powered to detect significant effects for these four comparisons at a sample anywhere above 2000 total respondents. The right panel of Figure 3 plots the minimum detectable effect size for H2 (difference between peer and non-peer cued stigma), which would logically be the more difficult treatment effect to identify, given that it compares primes that all include stigmatization attempts. Still, we see that our sample of 5,500 respondents will be well powered to detect effects as small as $d = 0.075$.

individual outcomes.

This gives us confidence that, should there be no discernible differences in the efficacy of treatments, we have estimated a tightly-bounded non-significant effect.

Figure 3: Power calculations



4 Conclusion

One of the most pressing political questions of recent times is how to tackle growing radical-right behavior across Western democracies. This paper looks into social norms as one avenue to fulfill that goal. In making radical-right behavior more socially costly, norms can deter individuals from engaging in it. This project validates experimental treatments and investigates the sources that are most effective in strengthening anti-radical-right norms. In doing so, it should offer three main contributions. First, we present easily generalizable primes of stigma and validate their efficacy. Validating these treatments and comparing their impacts will be useful for experimentalists interested in manipulating social norms surrounding the radical-right. Second, our findings will enhance our understanding of political norms. If source-specific cues of stigma perform differentially, this would imply that the source cuing stigma against the radical right is a key factor in how individuals process and update evaluations about norms. Finally, our findings have important policy implications. Extensive literature demonstrates the potent role of social norms in eliminating socially undesirable behaviors, including female genital cutting (Evans et al. 2019) or foot binding

(Mackie 1996). Our paper will highlight the extent to which similar interventions can also prevent undesirable behaviors in the political realm—namely radical-right behavior, which is typically exclusionary and threatening to minority groups (Mudde 2007).

While our study draws upon an experiment in Spain, our findings are likely to travel to other contexts. The radical right has been regarded as a party family with similar characteristics across Western democracies (e.g., Mudde 2007; Norris 2005). The same argument has been put forward when it comes to dynamics pertaining to the normalization and stigmatization of these parties (Valentim 2021; Ammassari 2024). As such, our results are likely to translate to cases beyond Spain.

References

Albornoz, Facundo, Jake Bradley, and Silvia Sonderegger. 2022. *Updating the social norm: the case of hate crime after the Brexit referendum* [in eng]. Working Paper 2022. Red Nacional de Investigadores en Economía. https://www.econstor.eu/handle/10419/228376.

Alvarez-Benjumea, Amalia, and Vicente Valentim. 2024. "The Enforcement of Political Norms: Theory and Evidence from Spain." *British Journal of Political Science,* https://doi.org/10.2139/ssrn.4271028.

Ammassari, Sofia. 2022. "It depends on personal networks: Feelings of stigmatisation among populist radical right party members." *European Journal of Political Research,* 1–19. https://doi.org/10.1111/1475-6765.12568.

———. 2024. "Deterrent or Stimulus? How Perceived Societal Stigma Affects Participation in Populist Radical Right Parties." *Political Studies,* 1–23. https://doi.org/10.1177/00323217241234527.

Axelsen, Jørgen. 2023. "The cordon sanitaire: a social norm-based model." *Journal of Elections, Public Opinion and Parties,* 1–21. https://doi.org/10.1080/17457289.2023.2168272.

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

Blinder, Scott, Robert Ford, and Elisabeth Ivarsflaten. 2013. "The Better Angels of Our Nature: How the Antiprejudice Norm Affects Policy and Party Preferences in Great Britain and Germany" [in en]. *American Journal of Political Science* 57 (4): 841–857. https://doi.org/10.1111/ajps.12030.

Bolet, Diane, and Florian Foos. 2023. "Media platforming and the normalisation of extreme right views," https://doi.org/10.31235/osf.io/urhxy.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin. 2020. "From Extreme to Mainstream: The Erosion of Social Norms." *American Economic Review* 110 (11): 3522–3548. https://doi.org/10.1257/aer.20171175.

Clayton, Katherine, Nicholas T. Davis, Brendan Nyhan, Ethan Porter, Timothy J. Ryan, and Thomas J. Wood. 2021. "Elite rhetoric can undermine democratic norms." *Proceedings of the National Academy of Sciences* 118 (23): e2024125118. https://doi.org/10.1073/pnas.2024125118.

De Jonge, Léonie, and Elizaveta Gaufman. 2022. "The normalisation of the far right in the Dutch media in the run-up to the 2021 general elections." *Discourse & Society* 33 (6): 773–787. https://doi.org/10.1177/09579265221095418.

Dinas, Elias, Sergi Martínez, and Vicente Valentim. 2024. "Social norm change, political symbols, and expression of stigmatized preferences." *The Journal of Politics* 86 (2): 488–506. https://doi.org/10.1086/726951.

Evans, W. Douglas, Cody Donahue, Jeremy Snider, Nafisa Bedri, Tibyaan A. Elhussein, and Samira Ahmed Elamin. 2019. "The Saleema initiative in Sudan to abandon female genital mutilation: Outcomes and dose response effects" [in en]. *PLOS ONE* 14 (3): e0213380. https://doi.org/10.1371/journal.pone.0213380.

Favero, Adrian, and Mattia Zulianello. 2023. "Building legitimacy: why the populist radical right engages in grassroots activism at the local level." *Journal of Contemporary European Studies* 0 (0): 1–12. https://doi.org/10.1080/14782804.2023.2269383.

FBI, National Press Office. 2023. *FBI Releases 2022 Crime in the Nation Statistics.* https://www.fbi.gov/news/press-releases/fbi-releases-2022-crime-in-the-nation-statistics.

Golder, Matt. 2016. "Far Right Parties in Europe." *Annual Review of Political Science* 19 (1): 477–497. https://doi.org/10.1146/annurev-polisci-042814-012441.

Gul, Sami. 2023. "Parliamentary representation of radical right and anti-immigration attitudes." *Electoral Studies* 86:102680. https://doi.org/10.1016/j.electstud.2023.102680.

Harteveld, Eelco, Stefan Dahlberg, Andrej Kokkonen, and Wouter van der Brug. 2019. "Social Stigma and Support for the Populist Radical Right: An Experimental Study" [in en]. *Scandinavian Political Studies* 42 (3-4): 296–307. https://doi.org/10.1111/1467-9477.12153.

Krause, Werner, Denis Cohen, and Tarik Abou-Chadi. 2023. "Does accommodation work? Mainstream party strategies and the success of radical right parties" [in en]. *Political Science Research and Methods* 11 (1): 172–179. https://doi.org/10.1017/psrm.2022.8.

Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification.* Harvard University Press.

Mackie, Gerry. 1996. "Ending Footbinding and Infibulation: A Convention Account." *American Sociological Review* 61 (6): 999–1017. https://doi.org/10.2307/2096305.

Mudde, Cas. 2007. *Populist radical right parties in Europe.* Cambridge, UK ; New York: Cambridge University Press.

———. 2016. "The Study of Populist Radical Right Parties: Towards a Fourth Wave." *C-REX Working Paper* 1:1–23.

Norris, Pippa. 2005. *Radical Right: Voters and Parties in the Electoral Market* [in en]. Cambridge University Press.

Oshri, Odelia, Liran Harsgor, Reut Itzkovitch-Malka, and Or Tuttnauer. 2023. "Risk Aversion and the Gender Gap in the Vote for Populist Radical Right Parties" [in en]. *American Journal of Political Science* 67 (3): 701–717. https://doi.org/10.1111/ajps.12696.

Romarri, Alessio. 2020. "Do Far Right Mayors Increase the Probability of Hate Crimes? Evidence From Italy" [in en]. *SSRN Scholarly Paper,* no. 3506811, https://doi.org/10.2139/ssrn.3506811.

Tankard, Margaret E., and Elizabeth Levy Paluck. 2016. "Norm Perception as a Vehicle for Social Change" [in en]. *Social Issues and Policy Review* 10 (1): 181–211. https://doi.org/10.1111/sipr.12022.

Valentim, Vicente. 2021. "Parliamentary Representation and the Normalization of Radical Right Support." *Comparative Political Studies* 54 (14): 2475–2511. https://doi.org/10.1177/0010414021997159.

———. 2024. "Political Stigma and Preference Falsification: Theory and Observational Evidence." *The Journal of Politics* 86 (4): 1382–1402. https://doi.org/10.1086/729966.

Van Heerden, Sjoerdje Charlotte, and Wouter van der Brug. 2017. "Demonisation and electoral support for populist radical right parties: A temporary effect" [in en]. *Electoral Studies* 47:36–45. https://doi.org/10.1016/j.electstud.2017.04.002.

Van Spanje, Joost, and Till Weber. 2019. "Does ostracism affect party support? Comparative lessons and experimental evidence." *Party Politics* 25 (6): 745–758. https://doi.org/10.1177/1354068817730722.

Appendix

Supplemental Information

Stigmatizing the radical right

Contents

A1 Questionnaire in Spanish 1
A1.1 Sociodemographics 1
A1.2 Pre-treatment political preferences 2
A1.3 Treatment arms 3
A1.4 Post-treatment questions 3
A2 Additional Results 5

A1 Questionnaire in Spanish

A1.1 Sociodemographics

1. ¿Es usted...?
    - Hombre
    - Mujer
    - No binario/otro
2. ¿Cuántos años cumplió en su último cumpleaños?
    - 18, 19, 20, ...
3. ¿Cuál es su nivel máximo de estudios terminados?
    - Menos de 5 años de escolarización
    - Educación primaria
    - Educación secundaria
    - Bachillerato
    - Formación profesional
    - Grado universitario o licenciatura
    - Máster
    - Doctorado
    - Otro
    - Prefiero no contestar
4. ¿Cuántas personas viven en su hogar?
    - --
5. Juntando los ingresos netos mensuales de todos los miembros de su hogar, ¿cuántos diría usted, aproximadamente, que son los ingresos mensuales de su hogar?
    - Sin ingresos
    - Menos o igual a 500€
    - De 501 a 900€
    - De 901 a 1200€
    - De 1201 a 1500€
    - De 1501 a 2500€
    - De 2501 a 4000€
    - De 4001 a 6000€

- Más de 6000€

6. ¿En cuál de estas situaciones laborales diría que se encuentra?
   - Empleado/a
   - Jubilado/a
   - En desempleo
   - Estudiante
   - Trabajo del hogar no remunerado
   - Otro
7. ¿En qué comunidad autónoma vive actualmente?
   - (List)
8. ¿Cuál es su código postal?
   - CP: _____

A1.2 Pre-treatment political preferences

9. Cuando se habla de política, se utilizan normalmente las expresiones izquierda y derecha. Situándonos en una escala que va del 0 al 10, en la que 0 significa "lo más a la izquierda" y 10 "lo más a la derecha," ¿en qué casilla se colocaría?
   - 0-10
10. ¿Podría indicar cómo de español se identifica en una escala de 0 a 10, donde 0 significa que no se siente nada español/a y 10 que se siente muy español/a?
   - 0-10

A1.3 Treatment arms

Table A1: Vignettes corresponding to each treatment arm in Spanish.

Control	**Peer Stigmatization**	**Media Stigmatization**	**Political Elite Stigmatization**
Vox es un partido político que se formó en 2013.	Vox es un partido que se considera **muy estigmatizado** por la ciudadanía de este país. Por ejemplo, si una persona cualquiera muestra públicamente, por la calle o en una plaza, su apoyo a Vox o alguna de sus propuestas, es altamente probable que otros vecinos o gente que pase por allí, tanto de derechas como de izquierdas, lo juzguen o menosprecien.	Vox es un partido que se considera **muy estigmatizado** por los medios de comunicación de este país. Por ejemplo, si algún/a periodista o comentarista político apoya públicamente a Vox o alguna de sus propuestas, es altamente probable que otros periodistas o comentaristas, tanto de derechas como de izquierdas, lo juzguen o menosprecien.	Vox es un partido que se considera **muy estigmatizado** por los políticos de este país. Por ejemplo, si algún político expresa públicamente su apoyo a Vox o alguna de sus propuestas, es altamente probable que otros políticos, tanto de derechas como de izquierdas, lo juzguen o menosprecien.

A1.4 Post-treatment questions

- **Factual manipulation check:** El texto anterior describía a Vox. ¿Diría que el texto discutió a Vox como un partido estigmatizado, o como un partido relativamente normal en España? (Estigmatizado; Ni estigmatizado ni normalizado; Normalizado)
- **First order normative evaluations:** En una escala de 0 a 10, ¿cómo de aceptable consideras que es apoyar a Vox?
- **Second order evaluations**
  - En una escala de 0 a 10, ¿Cómo de aceptable dirías que se considera a Vox entre *los ciudadanos españoles*?
  - En una escala de 0 a 10, ¿Cómo de aceptable dirías que se considera a Vox entre *los medios de comunicación españoles*?
  - En una escala de 0 a 10, ¿Cómo de aceptable dirías que se considera a Vox entre *los políticos españoles*?
- **Sanctions**
  - Imagine que está en una plaza cualquiera y ve a alguien con una camiseta con el logo de Vox. ¿Con qué probabilidad reaccionaría del siguiente modo? / Ahora

considere una persona cualquiera en esa misma situación. Imagine que esa persona está en una plaza cualquiera y ve a alguien con una camiseta con el logo de Vox. ¿Con qué probabilidad cree que esa persona reaccionaría del siguiente modo?

  * Gritaría[s] a esa persona
  * Insultaría[s] a esa persona
  * Intimidaría[s] físicamente a esa persona
  * Cuchichearía[s] sobre esa persona con un o una acompañante tuya
  * Negaría[s] la ayuda a esa persona en caso de emergencia
  * Evitaría[s] interactuar con esa persona

- **Voting intentions (see Footnote 2)**
  - Con respecto a Vox, ¿cómo describiría su probabilidad de votar a este partido en el futuro?
    * Nunca votaré a Vox
    * Es poco probable que alguna vez vote a Vox
    * No estoy seguro si alguna vez votaré a Vox
    * Probablemente vote a Vox en el futuro
    * Seguro que votaré a Vox en un futuro

A2 Additional Results

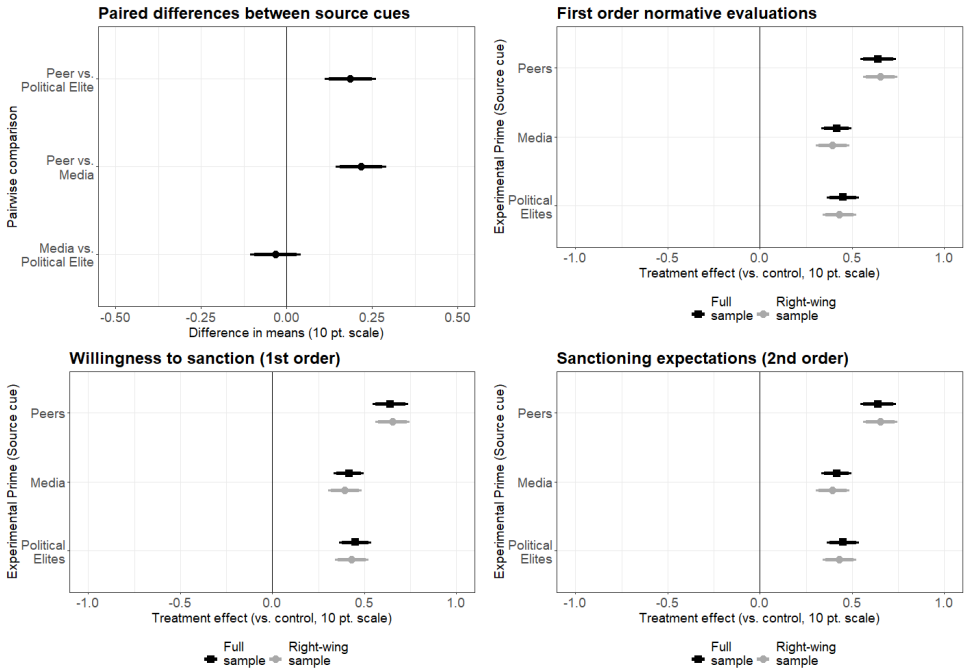



Figure A1: Pairwise comparisons; Simulated results on non-indexed outcomes