




LETTER

# Crossing the Line: Evidence for the Categorization Theory of Spatial Voting

Mark Pickup<sup>1</sup> , Erik O. Kimbrough<sup>2</sup>  and Eline A. de Rooij<sup>1</sup> 

<sup>1</sup>Department of Political Science, Simon Fraser University, Burnaby, BC, Canada and <sup>2</sup>Smith Institute for Political Economy and Philosophy, Argyros School of Business and Economics, Chapman University, Orange, CA, USA

**Corresponding author:** Mark Pickup; Email: [mark.pickup@gmail.com](mailto:mark.pickup@gmail.com)

(Received 3 February 2022; revised 1 March 2023; accepted 29 May 2023)

## Abstract

Bølstad and Dinas (2017) propose a model of spatial voting, based on social identity theory, that suggests supporting a candidate/policy on the other side of the ideological spectrum has a disutility that is not accounted for by common spatial models. Unfortunately, the data they use cannot speak directly to whether the disutility arises because individuals perceive their ideology as a social identity. We present the results of an experimental study that measures the norm against crossing the ideological spectrum; tests the cost of doing so, controlling for spatial effects; and demonstrates that this cost increases with the salience and strength of identity norms. By demonstrating the norm mechanism for the disutility of crossing the ideological spectrum, we provide strong support for B&D's model.

**Keywords:** norms; spatial voting; social identity; ideology

Ideology is often understood as a summary measure of an individual's policy preferences, projected onto the left-right spectrum (Downs 1957; Enelow and Hinich 1984). Evidence, though, suggests that ideology can also be the basis of social identity: people recognize those on their side of the spectrum as an in-group, and those on the other side as an out-group (Mason 2018); actions inconsistent with one's ideological identity generate negative emotions (Malka and Lelkes 2010); and individuals are willing to pay a cost to avoid these actions (Pickup, Kimbrough, and de Rooij 2022). This evidence further suggests that supporting a candidate/policy on the other side of the ideological spectrum has a disutility not accounted for by the proximity spatial model or the most common alternative, the directional spatial model (Bølstad and Dinas 2017).

Bølstad and Dinas (B&D) propose a *categorization theory of spatial voting* that accounts for this disutility, predicated on the idea that ideological identity is social identity. B&D use survey data from the European Election Study to evaluate a respondent's party preferences as a function of their ideological self-placement and their placement of the parties on a left-right spectrum. They observe a discontinuous change in respondents' party preference when comparing two parties on opposite sides of the political centre, which cannot be explained by directional or proximity voting.<sup>1</sup>

B&D's findings strongly corroborate their categorization theory of voting, but their use of observational data limits what they can test. One difficulty B&D faced arose from the fact that parties on

<sup>1</sup>They also examine 'critical cases' where the voter and party placements sharpen the differences in the predictions from proximity, directional and categorical voting models.

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

the same and opposite sides of the ideological spectrum as the respondent are clearly in different spatial positions. On the coarse ideological scale used, the difference between being on the same and opposite side of the ideological spectrum is at least 1/5 of its range (or more, depending on where the respondent places the parties). This confounds the effects of distance with the effects of crossing the centre. B&D provide compelling evidence of discontinuities in voters' party preferences at the centre of the ideological spectrum for a variety of voter self-placements and show that the effects of the centre on party preferences persist even when comparisons are restricted to parties at an equal ideological distance from the voter. Yet, B&D are unable to completely rule out the possibility that the discontinuity at the centre of the ideological spectrum is an artefact of some yet unknown distance function, as unlikely as that may seem.

Further, and more importantly, while their evidence is consistent with the categorization theory and highly persuasive, B&D cannot speak directly to whether this effect arises because individuals perceive their ideology as a social identity. We thus aim to augment their categorization theory by pinning down the mechanism through which the disutility of 'crossing the centre' arises: norms associated with an ideological identity. By their membership in an ideological identity group, people know that others expect them to follow a set of norms about which candidates/policies they should and should not support. They suffer social and psychological costs from violating norms and thus are disinclined to do so. Hence, the discontinuity is generated by the most basic ideological group norm – do not support the other side.

To assess the validity of this interpretation of categorization theory, our experimental study: measures the norm against crossing the ideological centre; tests the cost of doing so, controlling for spatial effects; and demonstrates that this cost increases with the salience and strength of identity norms. The resulting contribution extends beyond the literature on vote choice. By showing how norms shape individuals' assessments of those with an opposing ideological identity, we contribute to the literature on affective polarization (Mason 2018), which has suggested but not tested the idea that political identity norms contribute to polarization (Iyengar and Westwood 2015). Our findings also have potential implications for another voting theory – the reputational premium theory (Sniderman and Stiglitz 2012) – which we discuss in the concluding section.<sup>2</sup>

### A Categorization Theory of Voting

The categorization theory of spatial voting postulates that voters think in terms of ideological categories – left versus right, conservative versus liberal – and do so because an ideology can be a social identity (Bølstad and Dinas 2017). However, categorization theory does not offer an explicit account of how such identities influence behaviours and attitudes. We offer an interpretation of the categorization theory that builds on recent work arguing that identities matter for behaviour because of the *norms* associated with them (Groenendyk, Kimbrough, and Pickup 2022; Klar 2013; Pickup, Kimbrough, and de Rooij 2022). By norms, we mean shared expectations about how group members ought to behave by virtue of their group membership. We argue that a concern for following group norms influences decisions alongside personal preferences regarding the candidate/policy. This yields a model in which utility depends on both personal preferences and adherence to group norms, which is reminiscent of models of vote choice by Fiorina (1976), Jessee (2012), and Sniderman and Stiglitz (2012).<sup>3</sup>

Our model has two parts. The first captures the individual's *personal* preferences, including personal policy preferences. Here, the payoff an individual receives from supporting a particular

<sup>2</sup>Another alternative to the categorization theory is cognitive dissonance (Festinger 1957). It is possible that voting for a party on the other side of the ideological spectrum gives rise to disutility via cognitive dissonance. That said, our findings suggest the source of that cognitive dissonance – the violation of a norm that originates in a political identity. In other words, if it is norms that drive the categorization effect, this theory may not be so different from that of cognitive dissonance.

<sup>3</sup>Jessee's and Fiorina's models view departures from personal preferences as driven by partisanship, not ideology.

candidate/policy is a function of the distance between the position of their personal ideal candidate/policy and the position of the candidate/policy in some policy space. The model can accommodate many ways of defining this space (and the distance metric), which need not be restricted to a single dimension. For example, the proximity model uses the simplest metric -- the distance between a person's ideal candidate/policy and an actual candidate/policy on a left-right spectrum (Downs 1957). In evaluating alternatives, the payoff an individual receives from a candidate winning or a policy being implemented is weighted by the probability that the candidate/policy wins or is implemented (given that the individual supported the candidate/policy). The difference between this payoff and the payoff the individual receives from the candidate/policy *not* winning or *not* being implemented captures the *personal* considerations at stake when voting.

So far, the model is a classic spatial model with voters with well-defined preferences over policies/candidates (Downs 1957; Enelow and Hinich 1984). The model is augmented by assuming individuals also have preferences about norms, preferring (all else equal) to follow their identity groups' norms. We understand the influence of identities via the social and psychological consequences of experiencing or imagining the (dis)approval of other group members. The second part of the model thus considers the degree of disapproval group members expect to receive from violating a particular norm; this is the *norm strength* (Bicchieri 2006). However, a norm of a given strength may still affect two group members differently depending on the norm's *salience* (Ybarra and Trafimow 1998); that is, whether it is 'top of mind'. For example, political advertising may make the norms of one member's political identities more salient.

While normative considerations associated with any social identity may enter the voting model, our experiment exploits the intuition that ideological identity creates a discontinuity at the centre of the ideological spectrum, as implied by the categorization theory. If a candidate/policy is on the same side as the individual, support for that candidate/policy is likely consistent with ideological group norms (support those on 'your side'). If a candidate/policy is on the opposite side of the individual, support for that candidate/policy is likely inconsistent with ideological group norms (don't support the 'other side').<sup>4</sup> Our first contribution is to provide empirical verification of this intuition by measuring the strength of the norm against supporting policies/candidates from the other side of the ideological spectrum. We next demonstrate that the perceived cost of violating this norm affects reported vote preference, controlling for spatial considerations.

We then use another intuition from the model to demonstrate that the cost of violating ideological identity norms increases with the salience and strength of the norm. The intuition is that individuals can experience a conflict between their personal preferences and the norms of their ideological identity (or any politically relevant identity). For example, a small business owner may have preferences regarding the minimum wage that do not align with their liberal identity. We use response times to measure the conflict experienced by individuals when they express a personal preference inconsistent with the norms of their ideological identity. This conflict reflects the cost of violating the norms, and we use a priming manipulation to show that increasing the salience of such norms increases the conflict individuals experience. We also demonstrate that this effect is greater for those who perceive the norm as stronger.

### Experimental Design – Part 1

Our two-part experiment was included in a survey we commissioned from YouGov in Britain and the United States. Sample sizes are 3,219 and 3,000, respectively.<sup>5</sup> The first question (Q1) asks respondents to place themselves on a 0 to 100 ideological scale.<sup>6</sup> The remainder of our survey

<sup>4</sup>Of course, some policies are not unambiguously on a particular side; see our conclusion.

<sup>5</sup>Field dates and sampling procedures provided in Supplementary Information (SI) 1.

<sup>6</sup>Question wording provided in SI2.

excluded those that placed themselves in the centre of the spectrum (11 per cent of British and 12 per cent of American respondents) and those who did not identify themselves on the ideological spectrum (19 per cent of British and 10 per cent of American respondents).<sup>7</sup>

The second question (Q2) displays the respondent's self-placement on the ideological scale from Q1 and the position of two hypothetical parties: A and B. Party B is randomly placed either 0.5 units just left of centre or 0.5 units just right of centre. The randomization divides respondents into two groups. In the control group, party B is on the *same* side of the ideological spectrum as the respondent but only 0.5 units away from the centre. In the treatment group, party B is on the *opposite* side of the spectrum but only 0.5 units away from the centre. Party A is always on the same side as the respondent and just as far from the respondent as party B. When the respondent's self-placement is too close to one end for this to be possible, party A is at the most extreme position on the same side as the respondent. In this study, we focus on questions about party B, but the effect of the presence of party A is tested, and the results are even stronger without it (SI5).

The design allows us to measure and prime normative considerations against supporting a candidate from the other side. To demonstrate that such considerations exist, we need to demonstrate that a liberal (for example) believes she should vote for a liberal candidate, not just because of the spatial distance between the individual and candidate but because they both come from the same (liberal) group. This requires us to control for the spatial part of an ideological identity. Past work on this problem assumes a proximity spatial function (voters prefer the party/candidate spatially closest to them) and interprets any deviation from this as the effect of a non-spatial component (Fiorina 1976; Jessee 2012). The limitation is that if a different spatial function applies, this misspecification will be interpreted as evidence of a non-spatial component. Ultimately, a vast number of spatial functions may apply, including those with more than one spatial dimension.

Our design eliminates the need to know the functional form or the dimensionality of an individual's spatial ideology. We allow for the possibility that different individuals use different spatial functions. Consider the difference in Figure 1 in what is seen by a respondent ( $r_i$ ) depending on whether they are assigned to treatment or control.

Since we randomly assign subjects to treatment and control, we assume the average distance of respondents from party B is visually the same in the treatment and control groups, except that the centre line places party B in a different category for treatment group respondents but not for control group respondents.<sup>8</sup> Different responses to party B by those in treatment and control are then attributable to the categorization of the party (left/liberal or right/conservative) and not spatial differences.<sup>9</sup>

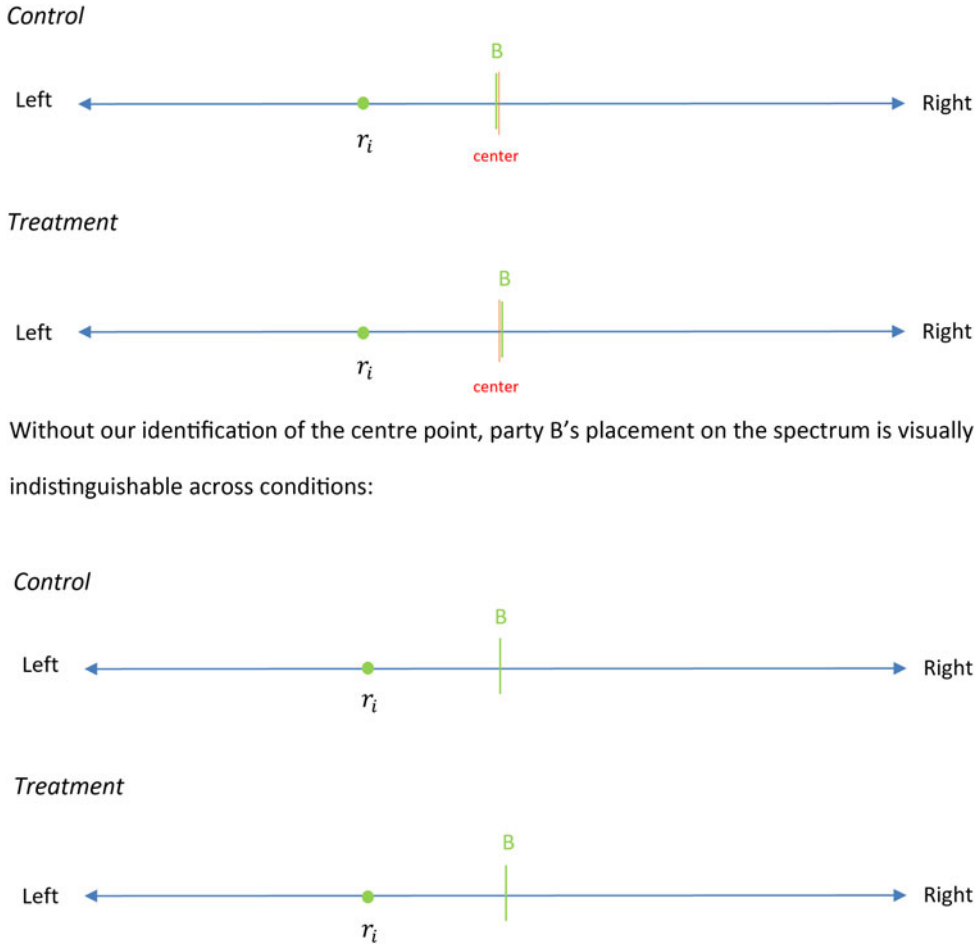
Q2 asks the respondent how much others with the same ideological identity would approve/disapprove of them voting for a candidate from party B (on a 0 to 10 scale). This captures the normative considerations derived from their ideological group about voting for party B. Since spatial differences are controlled, *the difference* between treatment and control is a direct measure of the ideological group norm against voting for the other side.

Q3 then asks how likely the respondent would be to vote for a candidate from party B (on a 0 to 100 scale). Again, because spatial differences are controlled, the difference between treatment and control is a direct measure of the disutility of violating the group norm that one should not vote for a party on the other side and its effect on vote choice.

<sup>7</sup>We exclude respondents that took less than five seconds to answer Q1 and less than ten seconds to answer Q2, on the basis that this is not long enough to read and understand the questions.

<sup>8</sup>For evidence of balance across groups on several pre-treatment covariates and distance from the centre, see SI3.

<sup>9</sup>A party on the other side might appear further because the centre acts as a heuristic for distance. This is consistent with the categorization theory if this is because the party is perceived to be in a different ideological category. The results we present about norms suggest that this is the case.



**Figure 1.** Experimental design. Treatment and control conditions, with centreline shown to respondents (upper figure) and without centreline (lower figure)

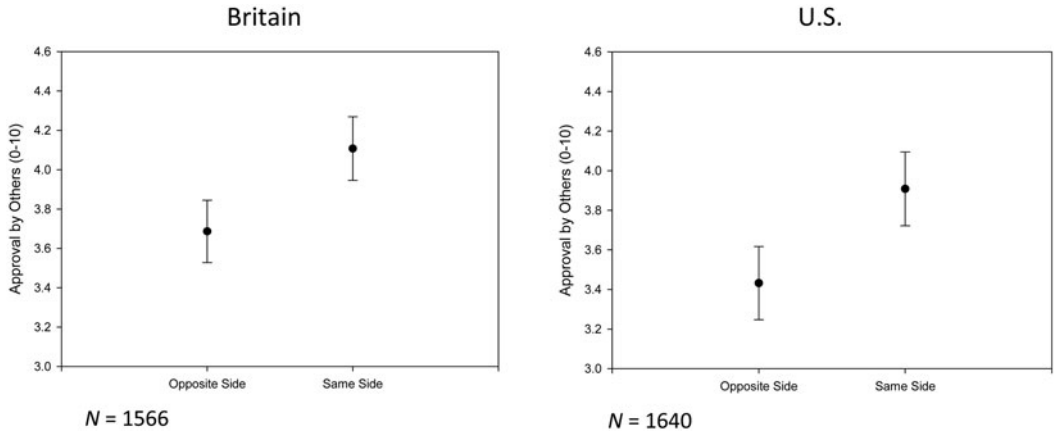
Hypothesis 1) *There is a norm against voting for the other side.* Controlling for spatial position, individuals expect others with the same ideological identity to approve more (disapprove less) of them voting for a candidate with the same ideological identity compared to a candidate with the opposite ideological identity.

Hypothesis 2) *Violating norms is costly and affects vote choice.* Controlling for spatial position, individuals are more likely to vote for a candidate with the same ideological identity compared to a candidate with the opposite ideological identity.

### Results – Part 1

Figure 2 shows the average expected approval of voting for a candidate from party B on the opposite and same side of the ideological spectrum. Consistent with our hypothesis, expected approval is higher if party B is on the same side. For British respondents, the difference is 0.42 (p-value < 0.001) on a 0 to 10 scale.<sup>10</sup> For American respondents, the difference is 0.48 (p-value < 0.001). The greater disapproval expected for voting for a candidate on the opposite

<sup>10</sup>All p-values are two-tailed.



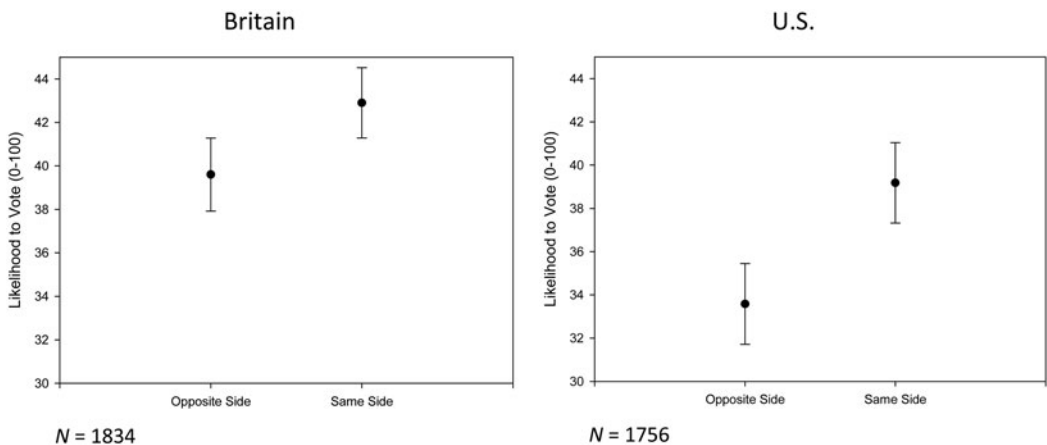
**Figure 2.** Effect of treatment on approval by others of voting for a candidate from party B, means with 95% confidence intervals

side vs the same side, holding spatial distance constant, demonstrates that individuals perceive a norm against crossing the centre of the ideological spectrum.<sup>11</sup>

To demonstrate that individuals anticipate a cost for violating this norm, we explore the effect of norms on vote utility by comparing the average likelihood of voting for a candidate from party B when that party is on the same side of the spectrum as the respondent and when that party is on the opposite side, again controlling for spatial differences. Looking at [Figure 3](#), we find that the likelihood of voting for a candidate from party B is higher when party B is on the same side. For British (American) respondents, the difference is 3.3 (5.6) on the 0 to 100 likelihood scale (p-values < 0.01).

### Experimental Design – Part 2

Given the evidence that voting for the other side is seen as a norm violation and creates a disutility that alters vote likelihood, we now aim to assess whether increasing the salience of ideological group norms increases this disutility. We do this by demonstrating that when respondents



**Figure 3.** Effect of treatment on likelihood to vote for a candidate from party B, means with 95% confidence intervals

<sup>11</sup>SI5 shows that this result survives robustness checks, such as controlling for the implied partisan identity of party B.

express a personal preference that is incongruent with the norms of their ideological group, they experience an internal conflict, and this conflict is increased by making the norms of their group more salient.

We use responses to a fourth question (Q4), which asks individuals their positions on 12 issues from a shortened version of the Wilson and Patterson conservatism scale (1968) developed by Henningham (1996). This question was only put to the British sample. Previous work shows that right-/left-wing individuals express broad normative agreement on what position members of their group should take on the issues from the conservatism scale (Pickup, Kimbrough, and de Rooij 2022). By comparing an individual's responses to the agreed-upon position as revealed in that study, we can identify incongruent responses: ones that violate group norms. We then use a response time test to determine if individuals experience a conflict when their responses violate group norms.

Prior evidence suggests that when reporting traits about the self that are incongruent with the normative expectations of the identity group, individuals take longer than when they are reporting traits that are congruent (Smith and Henry 1996). This is because the mental representation of the self and the group are linked as two dimensions of a single self-representation (Turner et al. 1987). Therefore, when an individual reports a trait incongruent with group norms, they experience a conflict. This conflict reflects the cost of violating group norms, and working through this conflict increases the time it takes to report that trait. Accordingly, we expect individuals with identity-incongruent preferences to take longer to report their positions than individuals without conflict.

Individuals who were asked about a candidate from a party on the opposite side of the spectrum (specifically Q2&Q3) were directly primed to think about the norm against voting for the other side. Those who were asked about a candidate on the same side of the spectrum still had ideology primed but did not have their group's norms explicitly primed. If the conflict an individual experiences when considering an incongruent personal preference is due to the cost of violating the norms of the identity, we would expect those primed to think about group norms to experience a stronger conflict and therefore take even longer to report incongruent positions. To identify evidence of a normative conflict, we regress the time it took respondents to respond to Q4 on the number of responses that were incongruent with group norms, interacted with whether the individual was primed to think of group norms:<sup>12</sup>

$$\text{response.time} = \beta_0 + \beta_1 \# \text{incongruent} + \beta_2 \text{norm.prime} + \beta_3 \# \text{incongruent} \\ \times \text{norm.prime} + \varepsilon$$

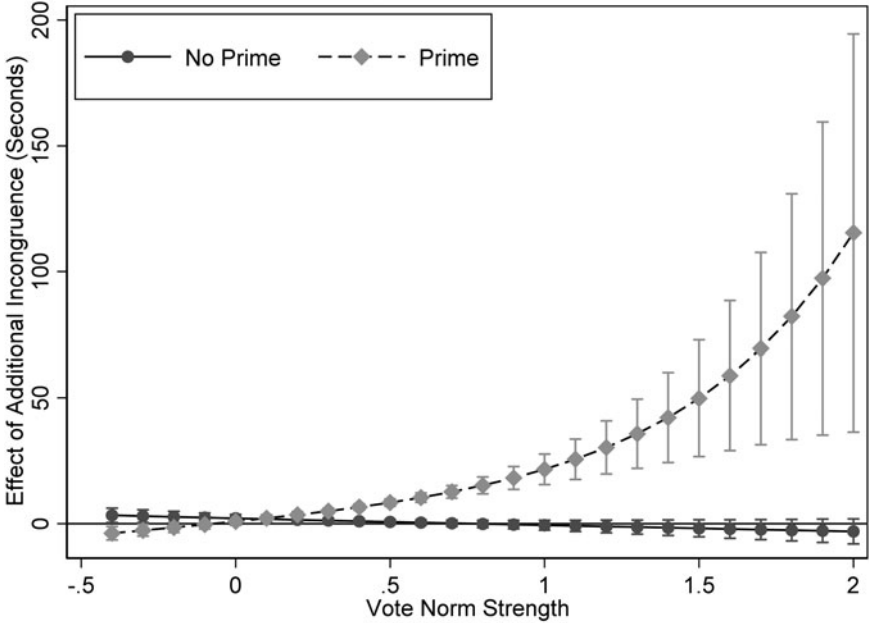
Hypothesis 3) *Norm salience increases norm violation costs.* When ideological group norms are made salient, individuals will take longer to report preferences that are incongruent with group norms.

## Results – Part 2

For each additional incongruent item, response time increases by an average of 0.83 seconds (p-value = 0.052) for those *not* primed to think about group norms and 7.05 seconds (p-value < 0.001) for those who were.<sup>13</sup> The norm prime is experimentally randomized, while the number of incongruent items is an observational variable, so it is the difference between these two effects that gets at the causal effect of norm priming. The difference is both large

<sup>12</sup>'Don't know' (DK) responses were treated as not incongruent, and the number of DKs was included as a control. The results are robust to treating DK as incongruent and not controlling for number of DK responses.

<sup>13</sup>SI6 for full model estimates.



*N* = 1881; 95% CI

**Figure 4.** Effect of incongruence by strength of vote norm (no norm prime and prime), maximum likelihood estimates from survival model with 95% confidence intervals

and statistically significant (*p*-value < 0.001) providing evidence that primed individuals experience greater conflict due to the increased cost of violating a group norm.

We should expect this effect of the prime to increase with an individual’s perception of the strength of group norms when expressing incongruent preferences.<sup>14</sup> We test this with a three-way interaction:

$$\begin{aligned}
 \text{response.time} = & \beta_0 + \beta_1 \# \text{incongruent} + \beta_2 \text{norm.prime} + \beta_3 \text{norm.strength} \\
 & + \beta_4 \# \text{incongruent} \times \text{norm.prime} + \beta_5 \# \text{incongruent} \times \text{norm.strength} \\
 & + \beta_6 \text{norm.prime} \times \text{norm.strength} + \beta_7 \# \text{incongruent} \times \text{norm.prime} \\
 & \times \text{norm.strength} + \varepsilon
 \end{aligned}$$

Figure 4 presents the effects of incongruence on response time. The effect increases with the magnitude of the vote norm for those primed to think of group norms. There is no such effect amongst those who were not primed. Norm strength is an observational variable, so again it is the difference between these two effects that identifies the causal effect of norm priming (*p*-value < 0.001).

**Discussion**

The categorization theory of voting introduced by B&D predicts that crossing the centrepoint of an ideological spectrum introduces a discontinuity in voters’ willingness to support a candidate.

<sup>14</sup>The strength of the vote norm (measured using Q2) is calculated for individuals at each point on the ideological spectrum. The norm is averaged over an interval of plus or minus 5 units on the ideological self-placement scale. (See SI4).



We argue that this discontinuity arises because voters with an ideological identity face a utility cost of violating their group's most basic norm: don't support the other side. We conduct a two-part experiment that provides a sharp test of the theory and of our proposed mechanism. We confirm that voters are averse to 'crossing the line' controlling for any spatial effects and that this aversion (at least in part) reflects group norms against supporting the other side. Our paper bolsters the important insights of B&D, highlighting a key mechanism of the categorization effect and using a controlled experiment to complement B&D's observational data.

Our findings have implications for another theory of vote choice: the *Reputational Premium* (RP) theory (Sniderman and Stiglitz 2012 (S&S)). This theory says that voters are particularly attracted to co-partisan candidates who take 'positions consistent with the policy outlook of their party' (p. 7). Set in the U.S. context, S&S propose the 'order rule' by which voters determine if a candidate is 'consistent': it is necessary and sufficient that the candidate is on the correct side of the other party's candidate. The order rule implies that a Republican voter gets an additional utility from supporting a Republican candidate so long as that candidate is to the right of the Democrat opponent. A Democrat voter gets an additional utility from supporting a Democrat candidate so long as that candidate is to the left of the Republican opponent.

At first glance, this looks similar to the categorization theory because both models assign a utility bonus to candidates that are on the 'correct side'; however, the models differ in whether they refer to the correct side of another candidate (RP theory) or to the correct side of the centre (categorization theory). They also differ in the mechanism by which this bonus occurs: a reputational premium in one case and norms in the other.

S&S give serious consideration to the hypothesis that the centre of the ideological spectrum matters. It is difficult to compare our empirics to theirs, given the different research designs (more on this in SI7) but S&S argue against this hypothesis on the grounds that there is no objective criterion for determining whether a candidate is on the right or left side of the policy space (S&S, p. 83). We accept that there is no objectively-defined centre or definition of liberal/conservative.<sup>15</sup> However, we disagree that informed voters (that is, programmatic partisans in S&S's language) are therefore unable (or unwilling) to label a candidate as liberal or conservative.

The categorization theory is about perceiving candidates/policies as falling within ideological categories, rather than objectively defining the categories, and the centre is simply the (perceived) boundary between categories. This can differ from voter to voter and across time and space; the voter need not even be able to articulate where the centre is in policy space. All that is required for the categorization theory to have implications for vote choice is that voters label candidates as liberal/left or conservative/right. To test whether voters can and will do this, we ran a survey 13–19 December 2022 ( $N = 1,000$ ). Amongst 'programmatic partisans' ( $N = 473$ ), 73 per cent identify their district's candidate for the U.S. House of Representatives in the previous midterm election (8 November 2022) as a 'liberal' or 'conservative'.<sup>16</sup>

This evidence demonstrates that many voters do place candidates into ideological categories, as the categorization theory requires. Coupled with our experimental evidence that the centre matters for vote choice, this has one of two implications for the relationship between the RP theory and categorization theory. The first possibility is that we might redefine where along the ideological spectrum a candidate can receive the RP. Our results suggest that being 'consistent with the policy outlook' might require being *labeled as* on the correct side of the ideological spectrum. I.e., this interpretation would suggest that the order rule is necessary but not sufficient to generate a utility premium because a candidate would need to be both on the correct side of the opposition *and* on the correct side of the spectrum. Alternatively, we might refine models of vote choice to include both an RP and a categorization effect, one of which is activated by being on the correct

<sup>15</sup>In our view, the definitions of ideological group categories are derived from the normative expectations of people who identify as such; see e.g., Groenendyk, Kimbrough, and Pickup (2022).

<sup>16</sup>The remainder said they did not know or that the candidate was 'other'. See SI8 for survey details.

side of an opposing candidate and the other of which is activated by being on the correct side of the ideological spectrum. Further experiments may be necessary to resolve this.

Another implication of our findings is that those who identify as moderate conservatives and moderate liberals may be less moveable than their moderate views would suggest, contributing to polarization. If voting for a candidate is perceived as supporting ‘the other side’, moderates may be unwilling to do so even if the other party candidate’s policy platform is closer to their ideal than the platform of the candidate on their own side.<sup>17</sup> Practically, this suggests that those seeking to build coalitions for policy change should avoid allowing their policies to become associated with an ideological label. The categorization effect may thus explain the appeal of political rhetoric identifying initiatives as bi-partisan or non-partisan; since parties are associated with an ideological ‘side’, rhetoric that divorces a policy from a party affiliation allows that policy to avoid being ‘tarnished’ in the eyes of some voters by its association with ‘the other side’ of the ideological spectrum.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0007123423000315>.

**Data availability statement.** Replication Data for this article can be found in Harvard Dataverse at: <https://doi.org/10.7910/DVN/X5AQO6>.

**Acknowledgements.** In addition to the journal reviewers and editors, we thank the audiences that attended presentations of earlier versions of this research at a variety of conferences and talks, including: the Nuffield College Politics Seminar; the Center for Comparative and International Studies (CIS) Colloquium at ETH Zürich; Simon Fraser University Economics Brown Bag; University of Southampton C2G2 Seminar; Stony Brook University Political Science Seminar; UC-Irvine Economics Seminar; University of Dayton Economics Seminar; University of Michigan SBEE seminar; the 2015 European Political Science Association Meeting; the 2016 Midwest Political Science Association Meeting; the 2016 Canadian Political Science Association Meeting; the 2016 Toronto Political Behavior Workshop; and the University of British Columbia Canadian-Comparative Politics Workshop. In particular, we thank Alan Jacobs, Scott Matthews, Edana Beauvais, Charles Breton, and Elias Dinas for their constructive comments. We are also very grateful to those at YouGov UK for their efforts in perfecting the instruments used in the experiment.

**Financial support.** This research was supported by the Social Sciences & Humanities Research Council of Canada, Grant #435-2016-1173.

**Competing interests.** The authors declare none.

**Ethical Standards.** The research was conducted in accordance with the protocols approved by the Simon Fraser University Office of Research Ethics.

## References

- Bicchieri C** (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bolstad J and Dinas E** (2017) A categorization theory of spatial voting: how the center divides the political space. *British Journal of Political Science* 47(4), 829–50.
- Downs A** (1957) *An Economic Theory of Democracy*. New York: Harper.
- Enelow JM and Hinich MJ** (1984) *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.
- Festinger L** (1957) *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Fiorina MP** (1976) The voting decision: instrumental and expressive. *Aspects Journal of Politics* 38(2), 390–413.
- Groenendyk E, Kimbrough EO and Pickup M** (2022) How norms shape the nature of belief systems in mass publics. *American Journal of Political Science*, 1–16. doi: 10.1111/ajps.12717
- Henningham JP** (1996) A 12-item scale of social conservatism. *Personality and Individual Differences* 20, 517–9.
- Iyengar S and Westwood S** (2015) Fear and loathing across party lines: new evidence on group polarization. *Am J of Pol Sci* 59(3), 690–707.

<sup>17</sup>This does raise the question of whether the ideological identity and its associated norms are as influential for a moderate as they are for a more extreme ideologue. Past research suggests that this may, in fact, not be the case (Pickup, Kimbrough, and de Rooij 2022).

- Jessee SA** (2012) *Ideology and Spatial Voting in American Elections*. New York: Cambridge University Press.
- Klar S** (2013) The influence of competing identity primes on political preferences. *Journal of Politics* 75(4), 1108–24.
- Malka A and Lelkes Y** (2010) More than ideology: conservative-liberal identity and receptivity to political cues. *Social Justice Research* 23, 156–88.
- Mason L** (2018) Ideologues without issues: the polarizing consequences of ideological identities. *Public Opinion Quarterly* 82 (S1), 280–301.
- Pickup M, Kimbrough EO and de Rooij EA** (2022) Expressive politics as (costly) norm following. *Political Behavior* 44(4), 1611–31.
- Pickup M, Kimbrough EO and de Rooij EA** (2023) Replication Data for: Crossing the Line: Evidence for the categorization theory of spatial voting. Available from <https://doi.org/10.7910/DVN/X5AQO6>, Harvard Dataverse, V1.
- Smith ER and Henry S** (1996) An in-group becomes part of the self: response time evidence. *Journal of Personality and Social Psychology* 22(6), 635–42.
- Sniderman PM and Stiglitz EH** (2012) *Reputational Premium*. Princeton: Princeton University Press.
- Turner JC et al.** (1987) *Rediscovering the Social Group: A Self-Categorization Theory*. Cambridge, MA: Basil Blackwell.
- Wilson GD and Patterson JR** (1968) A new measure of conservatism. *British Journal of Social and Clinical Psychology* 7, 264–69.
- Ybarra O and Trafimow D** (1998) How priming the collective self affects the relative weights of attitudes and subjective norms. *Personality and Social Psychology Bulletin* 24, 362–70.