Estimating the Between-Issue Variation in Party Elite Cue Effects

Ben M. Tappin¹*

¹Postdoctoral Scholar, Department of Political Science and Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, US

Abstract Party elite cues are among the most well-established influences on citizens' political opinions. Yet, there is substantial variation in effect sizes across studies, constraining the generalizability and theoretical development of party elite cues research. Understanding the causes of variation in party elite cue effects is thus a priority for advancing the field. In this paper, I estimate the variation in party elite cue effects that is caused simply by heterogeneity in the policy issues being examined, through a reanalysis of data from existing research combined with an original survey experiment comprising 34 contemporary American policy issues. My estimate of the between-issue variation in effects is substantively large, plausibly equal to somewhere between one-third and two-thirds the size of the between-*study* variation observed in the existing literature. This result has important implications for our understanding of party elite influence on public opinion and for the methodological practices of party elite cues research.

Party elites are among the most well-established influences on public opinion; many dozens of studies have documented evidence that exposure to party elite cues influences people's opinions about politics and policy. Broadly speaking, these studies find that, when people learn that in-party elites support (or oppose) a particular policy, they tend to become more supportive of (or more opposed to) the policy themselves, and vice versa when people learn the policy positions of out-party elites (Bullock 2011, 2020). However, despite the regularity of this phenomenon, there remain key gaps in our understanding of the influence of party elite cues on public opinion.

Chief among these gaps is the magnitude of their influence, because there exists large variation in the effects of party elite cues across studies. While

^{*}Corresponding author: Ben M. Tappin, Department of Political Science and Sloan School of Management, Massachusetts Institute of Technology, Sloan School of Management, 100 Main St, Cambridge, MA 02142, US. Email: benmtappin@gmail.com

[©] The Author(s) 2023. Published by Oxford University Press on behalf of American Association for Public Opinion Research. All rights reserved. For permissions, please email: journals.permissions@oup.com https://doi.org/10.1093/poq/nfac052

some studies estimate tiny effects on people's opinions, others estimate substantively large effects. For example, in an influential review, Bullock (2011) quantified the range in effect sizes across studies from a subset of the literature. He found that, across studies, the effect of party elite cues on people's opinions ranged from as small as 3 percent to as large as 43 percent of the scale used to measure opinions. In a more recent review, Bullock suggests this enormous variation has not diminished in the ensuing decade of research, concluding that "variation as great as this defeats most attempts to generalize about the sizes of party cue effects" (Bullock 2020, p. 6). This lack of generalizability is not benign (Findley, Kikuta, and Denly 2021; Linden and Hönekopp 2021). For example, it can entrench disagreement and stall progress on important research questions, such as the extent to which citizens' preferences follow (rather than lead) political elites (Zaller 1992; Adams et al. 2004; Lenz 2012; Sevenans 2021). Studies that observe large effects depict a citizenry whose opinions can be easily manipulated by elites; studies that observe small effects allow for the possibility that such manipulation is not so easy.

To enhance generalizability and facilitate theoretical progress, it is important to understand what explains the large variation across studies (Linden and Hönekopp 2021). There are many possible explanations because studies of party elite cue effects differ across a large number of dimensions, many of which plausibly influence the effect size of party elite cues. Perhaps most notably, existing studies often examine different policy issues, and typically just one or a small handful at a time (Guisinger and Saunders 2017). To illustrate this point, I recorded the number of policy issues examined in each paper across a sample of 50 published papers that reported a party elite cues study; the papers were obtained primarily via my knowledge of the literature and snowball sampling from the citation networks of existing party elite cue papers, and had a total citation count of 10,726 (for further details, see Supplementary Material section 1). Among this sample, the modal number of policy issues examined was 1–2; the median was 2; and the citationweighted mean was 2.65.

Heterogeneity in the policy issues examined across studies could explain the corresponding variation in party elite cue effect sizes for several reasons. For example, on some policy issues (but not others) people may already know where the parties stand prior to the study—they have been "pretreated." Similarly, on some issues (versus others) it may be easier for people to infer the positions of party elites from their prior knowledge of the parties' reputations (Gaines, Kuklinski, and Quirk 2007; Slothuus 2016). In either case, for studies examining such policy issues, the effect of exposure to the party elite cue is likely to be smaller, all else equal.

In addition to differences in pretreatment by the party cue, some issues are more personally salient to members of the public than others, or are more intensely covered by the media (Krosnick 1990; Ciuk and Yost 2016; Bullock 2020). Such high-salience issues are likely to be associated with stronger prior opinions than lower-salience issues. Thus, exposure to party elite cues on these issues is likely to have less of an impact on people's opinions, all else equal. Finally, policy issues vary in their general character (Carmines and Stimson 1980; Pollock, Lilie, and Vittes 1993; Mooney 2001). While some issues may be "easy" and rouse emotions by speaking to people's core values—such as policies related to abortion access—others may be "hard," highly technical, and disconnected from many people's core values—such as policies related to foreign affairs. It is plausible the latter types of issues are more amenable to influence by party elite cues because people are less motivated or able to think about such issues, and thus more likely to defer to party elites in a heuristic fashion (Kam 2005).

The foregoing is not meant to be an exhaustive list of the ways in which policy issues can differ. However, it suffices to suggest that there may exist substantial variation between policy issues in the size of party elite cue effects. Could this between-issue variation be substantial enough to explain the large variation in effects observed between *studies*? It is difficult to answer this question because, while existing studies examine different policy issues, they simultaneously differ over many other dimensions, many of which also plausibly influence the effect size of party elite cues.

To elaborate on just one example, existing studies often vary in the extent of baseline information that is available about the policy in question, ranging from scant background information-one or two sentences, or even nothing-through to detailed paragraphs of implications and arguments in favor of or against the policy. Insofar as the influence of party elite cues is attenuated when detailed policy information is available, as evidence suggests is the case (Boudreau and MacKenzie 2014; Tappin and McKay 2021), heterogeneity along this dimension plausibly explains variation between studies in party elite cue effects. In addition to heterogeneity in the availability of policy information, existing studies often also differ in whether the party elite cue is one-sided (in-party only) or two-sided (in-party and out-party), the strength of elite polarization, the source of the cue (e.g., party leaders or anonymous members of Congress), the method of measuring people's opinions, and the composition of the sample of respondents included in the study. Heterogeneity along all these dimensions, and more, plausibly explains variation between studies in party elite cue effects, and may not be independent from the particular policy issues examined.

In summary, despite an abundance of theories for why party elite cue effects might vary substantially across policy issues, there is a distinct lack of clear and systematic evidence regarding (1) the quantitative magnitude of this variation and (2) how the variation compares to—and therefore may explain-the large between-study variation observed in party elite cue effects.

In this paper, I aim to provide such evidence. To do so, I first reanalyze existing data from a recent study of party elite cue effects that included an unusually large number of policy issues (10) within the same experiment (Barber and Pope 2019). I selected this particular study for reanalysis because, to my knowledge, it reports a party elite cues experiment with the largest number of policy issues of any existing study.¹ It is important to maximize the number of policy issues because the smaller the sample of policy issues, the more uncertain the estimate of between-issue variation. Second, I conduct an original party elite cues experiment with a much larger sample of US policy issues—34 in total. Through my analysis of these two datasets, I estimate the variation in party elite cue effects that is due to heterogeneity in policy issues alone, with all the other aforementioned dimensions of heterogeneity held fixed by design. I am also able to explore what features of policy issues explain the variation.

To preview the main result, I find that variation across policy issues is substantial, plausibly equating to somewhere between one-third and twothirds the size of the between-*study* variation observed in the previous literature. This result suggests that a majority of existing between-study variation in party elite cue effects could be caused by existing studies' often small and idiosyncratic samples of policy issues, and has important implications for theory and practice in research on party elite influence. Second, I find some evidence that the variation in party elite cue effects across policy issues can be predicted by features of the issues; specifically, existing agreement with the in-party on the issue, and the general policy domain.

Reanalysis of Barber and Pope (2019)

Experiment Design

Barber and Pope (2019) conducted a 2017 survey experiment fielded by YouGov in which 1,300 US adults answered 10 questions related to public policy—indicating whether they supported or opposed each policy (or that they didn't know). The policy issues covered a variety of different domains, from a proposed increase in the minimum wage to the introduction of penalties against women who obtain abortions. The full list of questions is

^{1.} Broockman and Butler (2017) report an experiment with 17 policy issues, but their treatment is not a party elite cue treatment in the usual sense because the cue was from a state legislator and respondents were not told the party ID of the legislator. I also rule out reanalyzing Hill and Huber (2019) because their party elite cue treatment varies across policy issues, confounding variation across issues with that across treatments.

provided in Supplementary Material section 2.1. Before answering the questions, respondents were randomly assigned to one of four conditions, in which each policy question was accompanied by an endorsement from Donald Trump in a liberal direction (N = 200); an endorsement from Donald Trump in a conservative direction (N = 200); the position of Republicans in Congress (N = 400); or a control condition (no endorsements) (N = 500). Respondents were assigned to one and the same condition for all 10 questions, resulting in 13,000 observations. The authors' original outcome variable, *Support*, took a value of 1 if the respondent gave a liberal response and 0 if they gave a conservative response to the policy question.

Analytic Strategy

My goal is to quantify the variation in the size of party elite cue effects across policy issues and compare it to the between-study variation observed in the literature. Bullock (2011) provides an estimate of the latter in the form of the *range* in effect sizes observed across studies in the literature. He estimates a range of approximately 40 percent of the outcome scale (i.e., 43 percent to 3 percent). Thus, in my analyses I will estimate the range in effect sizes across policy issues, facilitating comparison with his estimate of 40 percent.

One strategy for computing the range is to use OLS to estimate the average treatment effect (ATE) of the party elite cue for each policy issue in the set, and then compute the distance between the minimum and maximum ATE across issues. However, this will tend to overestimate the range because some ATEs will be particularly large, and some particularly small, due to sampling variability alone—not due to true differences in treatment effects. That is, unless there is a large quantity of data for each individual policy issue, the raw estimates will tend to be noisy and overfit, making the range across policy issues appear larger than it actually is.

To more accurately estimate the range in average treatment effects (ATE) across policy issues, I use multilevel modelling (Gelman and Hill 2006; McElreath 2020). The informal description of this approach is that it provides an estimate of the variation in ATEs across policy issues, while explicitly accounting for the uncertainty with which each individual ATE is estimated. As a result, it provides a more accurate estimate of the true variation in ATEs across policy issues.² Using this estimate of the variation, I can compute the implied range in ATEs across policy issues, and compare it to

^{2.} This approach is similar to estimating the variation across ATEs (operationalized as the standard deviation) by fitting a random-effects meta-analysis to the raw estimates that were computed using OLS, as in, for example, Coppock, Hill, and Vavreck (2020).

Formally, I use the following multilevel linear regression specification:

$$Y_{jk} = Normal(\mu_{jk}, \sigma)$$
$$\mu_{jk} = \alpha_k + \lambda_j + \beta_k cue_{jk}$$
$$\begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \Sigma\right)$$
$$\Sigma = \begin{bmatrix} \sigma_{\alpha}^2 & \sigma_{\alpha}\sigma_{\beta}\rho \\ \sigma_{\alpha}\sigma_{\beta}\rho & \sigma_{\beta}^2 \end{bmatrix}$$
$$\lambda_i \sim Normal(0, \sigma_i)$$

where

 Y_{jk} is the outcome variable, assumed to be distributed normally, with mean μ_{jk} and standard deviation σ .

 α_k indexes the issue-specific random intercepts with mean α and standard deviation σ_{α} .

 λ_j indexes the respondent-specific random intercepts with mean of 0 and standard deviation σ_{λ} .

 β_k indexes the issue-specific random slopes on the average treatment effect of the party cue, with mean β and standard deviation σ_{β} .

The issue-specific random intercepts and random slopes are assumed to be distributed according to a multivariate normal distribution, where Σ is the variance-covariance matrix of the distribution.

Thus, the model allows the value of the intercept and average treatment effect to vary according to the particular policy issue in question. (A varying intercept allows for the fact that different policy issues may have different mean values of the outcome variable at baseline, that is, in the control group—the intercept is also allowed to vary across respondents.³) The variable *cue_{jk}* indicates whether the outcome variable was observed under treatment (1) or control (0). The parameter σ_{β} captures the variation in the average treatment effect across policy issues, in terms of its standard

^{3.} This accounts for the fact that in all analyses there are multiple observations per respondent, and respondents may differ in their average policy opinion at baseline.

deviation—this is the key quantity of interest that allows me to compute the range in treatment effects. I fit the model in a Bayesian framework,⁴ and specify weakly informative prior distributions on all model parameters (see Supplementary Material section 2.2 for the prior distributions, as well as tables of model results and diagnostics).

Finally, for all analyses, I must transform the outcome variable so that the expected sign of the party cue effect is in the same "direction" for all respondents and issues, otherwise computing the variation in the effect is nonsensical. To do so, I recode the policy opinions so that higher values indicate greater agreement with the in-party cue. For example, in the Barber and Pope (2019) data, if a Republican respondent gave a liberal response in the Liberal Trump versus control comparison, I code this as 1, whereas if a *Democrat* respondent gave a liberal response in the Liberal Trump versus control comparison in the Liberal Trump versus control comparison, I code it as 0. This recoding scheme reflects the theoretical expectation that cues from in-party elites cause updating toward the cue, while cues from out-party elites cause updating away from the cue (Nicholson 2012; Bullock 2020). I exclude pure Independents from the analysis but include party "leaners," and the analysis is unweighted.

Results

Figure 1 shows the results of my reanalysis of Barber and Pope (2019).⁵ There are three sets of results, corresponding to one model for each of the three treatment-control comparisons (Liberal Trump, Conservative Trump, or Republicans in Congress). Each row of figure 1 corresponds to one set of results. The left-hand panel in each row shows the raw estimates of the average treatment effect (ATE) of the party elite cue for each policy issue, computed by OLS. The dark vertical line shows the mean ATE, estimated by the multilevel model across all 10 policy issues; that is, parameter β in the multivariate normal distribution above. The lighter shaded region is the 95 percent highest posterior density interval (HPDI), which is the narrowest region that covers the value of the parameter with 95 percent probability, given the data and model.

So, for example, the mean ATE for the Liberal Trump cue treatment is 0.09; that is, 9 percent of the outcome scale. However, there is of course variation across policy issues around this mean ATE, and the variation is the quantity of interest. For instance, the policy question with the largest ATE in

^{4.} Using the R package *brms* (Bürkner 2017).

^{5.} The results in this section differ from those reported in Barber and Pope (2019) because those authors reported models separated by party ID. The results of my models separated by party ID, reported in Supplementary Material section 2.3, closely reproduce theirs.



Figure 1. Reanalysis of Barber and Pope (2019). (Panels A, C, E) Raw average treatment effect (ATE) of party elite cue on policy support for each policy issue. The mean ATE across policy issues is shown by the solid vertical line (the 95 percent credible interval is shown by the shaded area). Note that the raw estimate for Immigration is missing from panel C because of a technical error in the original survey conducted by Barber and Pope (2019). (Panels B, D, F) Distribution and range of ATEs estimated by the multilevel models. The 95 percent credible intervals are shown in brackets.

response to the Liberal Trump cue was *Climate Change* (0.16) (figure 1A). The policy question with the smallest ATE was *Guns Background* (0.06). In contrast, for the Conservative Trump cue, *Guns Background* had the largest ATE, and *Climate Change* had the smallest (figure 1C).

As described above, naively interpreting the range of the raw estimates across policy issues can give an inaccurate picture of the variation because of sampling variability. That is why I fitted the multilevel model. The model captures the variation in terms of the standard deviation of ATEs across policy issues, which it estimates to be 0.01 (Liberal Trump), 0.06 (Conservative Trump), and 0.03 (Republicans in Congress) across the three types of cue treatment, annotated in the right-hand panels of figure 1.

Together, the mean ATE and the standard deviation describe a (normal) distribution of ATEs across the unobserved population of policy issues, estimated from the data. The estimated distributions are plotted in the right-hand panels of figure 1, with the raw estimates overlaid as a histogram. To compute the expected range of ATEs across policy issues, I simply take the mean \pm 2.3 standard deviations, which covers 98 percent of the distribution. Thus, if variation in party elite cue effects were due to policy issues alone, we would expect the large majority of party elite cue effects (49 out of 50) to fall within the estimated range.⁶ The estimates of the range are shown at the base of each distribution in the right-hand panels of figure 1, with 95 percent HPDIs. For the three different cue treatments, the range is 0.06 (Liberal Trump), 0.26 (Conservative Trump), and 0.14 (Republicans in Congress), or 6 percent, 26 percent, and 14 percent of the outcome scale, respectively. On average, this is over one-third the size of the 40 percent range observed across previous studies in the literature (Bullock 2011). Therefore, this analysis suggests that a large minority of existing between-study variation in party elite cue effects could potentially be explained by existing studies' often small and idiosyncratic samples of policy issues.

While this reanalysis offers insight into the extent of variation in party elite cue effects that may be due to policy issues alone, it is limited by the small number of policy issues included in the original experiment. In particular, there remains much uncertainty about the magnitude of the range. Thus, I conducted a new party elite cues survey experiment with a much larger sample of US policy issues, described next.

Original Survey Experiment

Policy Issues and Party Elite Cues

To source a large and diverse sample of US policy issues, and corresponding party elite cues, I used the website www.isidewith.com, an online political

^{6.} I compute the range to cover 98 percent of the distribution rather than, say, 95 percent, because there are many dozens of party elite cue studies in the existing literature—more than 50 in total. Therefore, a range with 98 percent coverage provides a more useful benchmark for how many existing party elite cue effects we would expect to fall within the range *if* the variation between studies was due to policy issues alone (i.e., 49 in 50).

encyclopedia that documents the policy positions of US political candidates and leaders. I searched the database and recorded the policy positions of then-president Donald Trump and former president Barack Obama, whom I judged to be the two most prominent Republican and Democratic party elites at the time of the experiment (August 2020).

My initial search yielded 74 distinct policy questions for which the two elites had positions as documented by the website. In my initial search, I included only those policy questions on which the listed source for the elites' position was their "public statements" or "voting record," excluding those policy questions for which the source was simply listed as the elites' "voter base." Of the 74 policy questions in the initial search set, Trump and Obama ostensibly adopted the same position on 37 of them. This left 37 policy questions on which the two elites disagreed, and it is these questions that I selected for use in the experiment. Finally, of these 37 selected policy questions, I dropped two which referred to issues that were no longer of relevance—for example, the question "Should the U.S. send ground troops into Syria to fight ISIS?"—and one on which the listed positions of Trump and Obama were not in clear disagreement. This left 34 policy questions for my experiment, with corresponding Trump and Obama cues. The policy questions covered a wide diversity of issues (table 1).

Data

The survey was fielded in August 2020 to 1,730 US adults recruited via Lucid, a marketplace for online survey research that uses quota sampling to match respondents to the US Census demographics of age, gender, education, and region (Coppock and McClellan 2019). While samples recruited via Lucid are not fully representative of the general US population, a growing body of evidence suggests this does not much compromise the generalizability of my results: formal comparisons between treatment effects estimated using convenience samples, versus those in national samples, indicate the two often correspond closely (Mullinix et al. 2015; Coppock 2019). For information about sample characteristics, see Supplementary Material section 3.1.3.

Experiment Design

Survey respondents were shown a random subset of six of these policy questions, in a randomized order, and were asked to give their opinion on each question before moving on to the next. All opinions were given on 1-7 Likert scales, anchored with respect to the question that was posed. For example, the large majority of the questions were yes/no answer format and the corresponding scale anchors were "Definitely Not," "No," "Probably No,"

Policy ID	Policy question	Obama	Trump
1	Should convicted criminals have the right to vote?	Yes	No
2	Should prisons ban the use of solitary confinement for juveniles?	Yes	No
3	Should police officers be required to wear body cameras?	Yes	No
4	Should victims of gun violence be allowed to sue firearms dealers and manufacturers?	Yes	No
5	Should it be illegal to burn the American flag?	No	Yes
6	Should local police increase surveillance and patrol of Muslim neighborhoods?	No	Yes
7	Should an in-state sales tax apply to online purchases of in-state buyers from out-of-state sellers?	Yes	No
8	Should the government require businesses to pay sal- aried employees, making up to \$46k/year, time-and- a-half for overtime hours?	Yes	No
9	Should the U.S. raise taxes on the rich?	Yes	No
10	Should the government prevent "mega mergers" of corporations that could potentially control a large percentage of market share within its industry?	Yes	No
11	Should pension payments be increased for retired government workers?	Yes	No
12	Should welfare recipients be tested for drugs?	No	Yes
13	Should the current estate tax rate be decreased?	No	Yes
14	Should the President offer tax breaks to individual companies to keep jobs in the U.S.?	No	Yes
15	Should the Federal Reserve Bank be audited by Congress?	No	Yes
16	Should the U.S. continue to participate in the North American Free Trade Agreement (NAFTA)?	Yes	No
17	Should political candidates be required to release their recent tax returns to the public?	Yes	No
18	Should corporations, unions, and non-profit organiza- tions be allowed to donate to political parties?	No	Yes

Table 1. Policy questions and party elite cues.

(continued)

Policy ID	Policy question	Obama	Trump
19	Should there be a 5-year ban on White House and Congressional officials from becoming lobbyists after they leave the government?	No	Yes
20	Should disposable products (such as plastic cups, plates, and cutlery) that contain less than 50 percent of biodegradable material be banned?	Yes	No
21	Should the U.S. withdraw from the Paris Climate Agreement?	No	Yes
22	Should the U.S. increase or decrease foreign aid spending?	Increase	Decrease
23	Should the U.S. remain in NATO?	Yes	No
24	Should the government cancel production of the F-35 fighter?	Yes	No
25	Do you support a single-payer healthcare system?	Yes	No
26	Should there be more or less privatization of veter- ans' healthcare?	Less	More
27	Should the U.S. build a wall along the southern border?	No	Yes
28	Should the U.S. increase or decrease the amount of temporary work visas given to high-skilled immigrant workers?	Increase	Decrease
29	Should the government require children to be vacci- nated for preventable diseases?	Yes	No
30	Should producers be required to label genetically engineered foods (GMOs)?	Yes	No
31	Should the government continue to fund Planned Parenthood?	Yes	No
32	Should "gender identity" be added to anti-discrimina- tion laws?	Yes	No
33	Should gay couples have the same adoption rights as straight couples?	Yes	No
34	What is your stance on abortion?	Pro-choice	Pro-life

Table 1. Continued.

"Not sure/Need more information," "Probably Yes," "Yes," and "Definitely Yes." On each question, respondents were randomly assigned to either the treatment group, in which they learned the positions of both Trump and Obama, or the control group, in which no such information was given (Supplementary Material section 3.1.2 illustrates how the treatment appeared to respondents). Randomization to treatment or control was independent across policy questions, meaning that the same respondent could be in the control group for some questions and the treatment group for other questions. Some readers might be concerned that this design means that party elite cues on the first question would prime respondents to think about the party's preferred position on subsequent questions-even when no party elite cue was provided-thus diminishing the apparent effect of the cue. However, in Supplementary Material section 3.2.2, I show that the data do not bear this concern out: the average treatment effect of the party elite cue, as well as its estimated variability across policy questions, is similar irrespective of the order in which policy questions were answered. Randomization was successful and there was negligible post-treatment attrition (see Supplementary Material sections 3.1.4 and 3.1.5).

Recent work indicates that the risk of demand effects biasing treatment effects is low in designs like the one I use here (Mummolo and Peterson 2019; Clifford, Sheagley, and Piston 2021). Nevertheless, to avoid demand effects I supplemented the 34 policy questions in table 1 with 11 questions on which Trump and Obama took the *same* position (not shown in table 1; see Supplementary Material section 3.1.1). Thus, there were 45 policy questions in total, of which eight were drawn randomly to be shown to each respondent; meaning that, in expectation, respondents would see approximately six of the "target" 34 policy questions (i.e., $34/45 \approx 0.75$ and 6/8 = 0.75). The 11 supplemental questions are not included in my analyses, serving merely as foils in the experimental procedure.

Analytic Strategy

I use the same model specification as in the reanalysis of Barber and Pope (2019). As before, I transform the outcome variable so that higher values indicate greater agreement with the in-party cue. For example, a Democrat respondent who answered "Definitely Yes"—originally a score of 7 on the opinion scale—to the fifth policy question in table 1 (illegal to burn US flag) would receive a score of 1, since the in-party cue (Obama's position) is to oppose this policy. Following this recoding scheme, all answers to the policy questions are scaled to lie between 0 and 1 so that treatment effects can be interpreted as percentage movement on the opinion scale—enabling comparison with the estimated range from Bullock (2011). Pure Independents are not included in the recoding scheme nor in the analysis (but party "leaners" are

included), leaving 1,240 respondents corresponding to a total of 7,460 observations for analysis.

Recent evidence suggests that rates of attention declined markedly among respondents recruited via Lucid through 2020-when the current data were collected-with corresponding decreases in the quality of provided data (Ternovski and Orr 2022). Thus, as a robustness check I fit an additional model, subsetting the data to include only those who passed a stringent pretreatment attention check in the survey. The attention check showed an image of the word "hello." Respondents were asked to write the letters in the image in alphabetical order (ehllo). Of the 1,240 respondents included in the full-sample analysis, 384 (\sim 31 percent) passed this attention check, resulting in a total of 2,297 observations for the model fit with attentive respondents. demographics of attentive respondents The the are reported in Supplementary Material section 3.1.3.

Results

The presentation format of the results follows that of the previous analysis. Figure 2 shows the results of the multilevel model fitted on the full sample of respondents. (A table of model results and diagnostics is in Supplementary Material section 3.2.1.) Figure 2A shows that the mean average treatment effect (ATE) across the 34 policy issues is estimated to be 0.08, or 8 percent of the outcome scale. The policy question with the largest ATE was number 15, "Should the Federal Reserve Bank be audited by Congress?" (raw estimate = 0.21). The question with the smallest ATE was number 3, "Should police officers be required to wear body cameras?" (-0.06).

The standard deviation of ATEs across policy issues is estimated as 0.03 (figure 2B). Thus, the computed range is 0.16 (the mean \pm 2.3 standard deviations); that is, 16 percent of the outcome scale.⁷ Similar to the previous analysis, this is over one-third the size of the 40 percent range observed across existing studies in the literature (Bullock 2011).

Figure 3 shows the results of the model fitted on the attentive respondents. While the estimate of the mean ATE is similar to the full sample, the standard deviation in the ATE notably increases in size, to 0.06. As a result, the range in ATEs across policy issues is 0.27, or 27 percent of the outcome scale among attentive respondents, approximately *two-thirds* the size of the 40 percent range observed across previous studies in the literature. Thus, together, the results from the original survey experiment suggest that the range in the ATE across policy issues is somewhere between one-third (16 percent)

^{7.} In Supplementary Material section 3.2.2, I show that the mean ATE and standard deviation of the ATE across policy issues are similar irrespective of the order in which the policy questions were answered by respondents (i.e., there is no evidence of order effects).



Figure 2. Analysis of original survey experiment (total N = 1,240 respondents; 7,460 observations). (Panel A) Raw average treatment effect (ATE) of party elite cue on policy support for each policy issue (for a given policy issue, minimum N = 190 respondents). The mean ATE across policy issues is shown by the solid vertical line (the 95 percent credible interval is shown by the shaded area). (Panel B) Distribution and range of ATEs estimated by the multilevel model. The 95 percent credible intervals are shown in brackets.

Between-Issue Variation in Cue Effects



Figure 3. Analysis of original survey experiment: Attentive respondents only (total N = 384 respondents; 2,297 observations). (Panel A) Raw average treatment effect (ATE) of party elite cue on policy support for each policy issue (for a given policy issue, minimum N = 47 respondents). The mean ATE across policy issues is shown by the solid vertical line (the 95 percent credible interval is shown by the shaded area). (Panel B) Distribution and range of ATEs estimated by the multilevel model. The 95 percent credible intervals are shown in brackets.

B.M. Tappin

16

and two-thirds (27 percent) the size of the between-study range observed in the literature.

Notably, these estimates cohere with two out of three estimates of the range from the reanalysis of Barber and Pope (2019), 14 percent and 26 percent (see figure 1). Taking the results of these two sets of analyses together therefore suggests that a majority of existing between-study variation in party elite cue effects could potentially be explained by existing studies' often small and idiosyncratic samples of policy issues.

Given the substantial variation in the effects of party elite cues across policy issues, it is natural to wonder what explains this variation. I discussed several possibilities in the introduction, including the likelihood of "pretreatment," the public salience of the issue, and the broad domain or character of the issue. Though not the primary focus of this paper—meaning I am unable to causally distinguish these factors with the data at hand—I *am* able to offer some relevant and new descriptive evidence, on two counts.

First, as well as estimating the standard deviation in the intercept and treatment effect across policy issues, the multilevel model also estimates the *correlation* between these parameters across policy issues (ρ in the covariance matrix). Since the intercept captures the average policy opinion in the control group, it can be interpreted as the extent of agreement with the in-party position at baseline—that is, absent exposure to the party cue. Thus, the correlation between the intercept and treatment effect across policy issues tells us whether existing agreement with the in-party on an issue predicts the influence of exposure to the party elite cue. Inspecting this correlation from the original survey experiment shows that it is strongly negative, and reliably different from zero, for both the model fitted on all respondents and the model fitted on attentive respondents: -0.70, 95 percent HPDI [-0.95, -0.28], and -0.77 [-0.99, -0.35], respectively.

The effect of the party elite cue was thus weaker for policy issues that had higher agreement with the in-party at baseline. This could arise because such issues are more likely to have been "pretreated" by the party cue before the experiment, and thereby exhibit smaller effects of (further) exposure to the cue during the experiment (Gaines, Kuklinski, and Quirk 2007; Slothuus 2016). Accordingly, this negative correlation coheres with one theory for why there is variation in party cue effects across policy issues (although there are other possible explanations of the correlation).

As a second exploratory analysis, I examine whether there are differences in party elite cue effects across policy *domains*. If particular policy issues are more amenable to party elite influence owing to their general character or domain, we should expect evidence of differences in party elite cue effects across broad policy domains.

To explore this idea, I consult the policy domain categories assigned to each of the 34 individual policy issues in the original survey experiment by the analysts at www.isidewith.com. There are 10 such domains, including foreign policy, healthcare, the economy, and social policy (Supplementary Material section 3.3.1 reports which domains are assigned to which issues).⁸ I proceed to fit a new multilevel model, updating the specification to include another dimension along which the average treatment effect of the party elite cue is allowed to vary: the broad policy domain. All other model specification details are the same as the initial analysis of the original survey experiment. (Supplementary Material section 3.3 reports the model specification, results, and diagnostics in full.)

Figure 4 shows the model-estimated average treatment effects (ATE) for the model fitted on all respondents (figure 4A) and attentive respondents (figure 4B). The estimates suggest nontrivial differences in the ATE between domains. In both models, the ATE for foreign policy issues is consistently among the strongest (10–11 percent of the outcome scale), whereas the ATEs for crime and social policy issues are consistently among the weakest (4–7 percent). This somewhat aligns with the classic work of Carmines and Stimson (1980), which advanced a foreign policy issue (the Vietnam War) as an example of a "hard" issue. If such issues are generally harder for people to engage with, it follows that people may tend to defer more to party elite cues on these issues.

Nevertheless, the rank-orderings in figure 4 should be held lightly, for several reasons. First, it is somewhat subjective which policy issues belong to which domains, and different categorization schemes could change the rank order. Second, with a sample of only 34 policy issues, none of the domain categories here contain a large number of individual policy issues—meaning there is not much data with which to confidently determine differences between domains if they exist. Future work can build on these initial estimates of domain-general differences in party elite cue effects.

Conclusion

It has long been recognized that between-study variation in party elite cue effects is substantial, undermining the generalizability and theoretical progress of party elite cues research (Bullock 2011, 2020). Understanding the causes of this variation thus constitutes a priority for advancing the field (Findley, Kikuta, and Denly 2021; Linden and Hönekopp 2021). In this paper, I estimated large variation across policy issues in the size of party elite cue effects, plausibly equal to somewhere between one-third and two-thirds the size of the between-*study* variation observed in the existing literature. This advances research on the influence of party elite cues in several ways.

8. This analysis is post-hoc exploration. Thus, I elected not to categorize the policy issues into domains myself because I am not blind to the results for individual policy issues and this may have inadvertently biased my categorization decisions.



Figure 4. Analysis of original survey experiment: Average treatment effect (ATE) of party elite cue for each policy issue domain. The mean ATE across policy issues is shown by the solid vertical line (the 95 percent credible interval is shown by the shaded area). (Panel A) Among all respondents (N = 1,240; 7,460 observations). (Panel B) Among attentive respondents (N = 384; 2,297 observations).

vnloaded from https://academic.oup.com/poq/advance-article/doi/10.1093/poq/nfac052/7017587 by guest on 31 January 2

Improving Generalizability

First, while my results do not *prove* that a large share of the variation between existing studies is due to heterogeneity in policy issues alone, they reveal it as a distinct possibility given that the typical study examines just one or a small handful of policy issues (e.g., see Supplementary Material section 1). What my results do clearly show, however, is that scholars of party elite influence should strive to expand the samples of policy issues that they examine in their studies, as this is likely to significantly increase the generalizability of their estimates (see also Clifford, Leeper, and Rainey 2021).

Importantly, this enhanced generalizability would accrue not only to estimates of party elite cue effects themselves, but also to additional quantities that are of interest to a wide range of scholars of party elite influence. For example, scholars are interested in the extent to which elite influence diminishes in the presence of other types of information (Boudreau and MacKenzie 2014; Agadjanian 2020), whether elite influence is larger than that of policy information (Cohen 2003; Bullock 2011), the mechanisms through which elite cues exert their influence (Petersen et al. 2013; Ehret, Van Boven, and Sherman 2018; Van Boven, Ehret, and Sherman 2018), how long their influence persists (Tappin and Hewitt 2021), which types of individuals are most susceptible (Bakker, Lelkes, and Malka 2020), and crosscultural comparisons of these and other related quantities (Brader and Tucker 2012). In many such cases, scholars would like to generalize their results beyond the specific policy issue(s) in their study.

However, if any of the aforementioned quantities depend in any way upon the magnitude of party elite influence—or other features of policy issues then they also depend upon the specific policy issue(s) being examined. Thus, the generalizability of those quantities stands to benefit from an expanded sample of policy issues. For example, Tappin and McKay (2021) studied whether the influence of party cues is diminished when people have access to substantive policy information, incorporating many more policy issues into their study design (20) than previous work. They found that, for some policy issues, the influence of the party cue diminished by approximately 100 percent of its original size when people had access to policy information; while, for other issues, the reduction in the influence of the party cue was much more muted—as little as 25 percent of its original size. It is easy to imagine different studies arriving at different conclusions regarding the extent to which exposure to policy information attenuates the influence of party cues, all because they studied a different handful of policy issues.

This illustrates the benefits to knowledge of expanding the sample of policy issues in studies of party elite influence; it avoids a published literature that consists of many studies canvassing a small number of policy issues that reach different, perhaps mutually exclusive, conclusions due to an idiosyncratic sample of issues. Such a literature can foment disagreement and stall theoretical progress (Yarkoni 2020; Linden and Hönekopp 2021). Importantly, as my experiment design and analytic approach demonstrate, expanding the sample of policy issues does not demand a large increase in the number of respondents (and thus study cost). Rather, it simply means randomizing respondents to a (larger) number of policy issues. While this may be infeasible for some study designs, it is eminently feasible for many survey and even field experiments—the design and analyses reported here offer a concrete template.

In a related sense, my results bear on the long-standing question over the extent to which citizens' preferences follow (rather than lead) political elites (Zaller 1992; Adams et al. 2004; Lenz 2012; Sevenans 2021). Studies that find large effects of party elite cues suggest that citizens' opinions may often be easily manipulated by political elites, while studies that find small or null effects allow for the possibility that citizens are at least somewhat circumspect in their opinion formation. My results illustrate that the drawing of general conclusions regarding this question from studies focused on a typically small sample of policy issues is unwise: effects may be strikingly large for some issues, consistent with elite manipulation, and small or non-existent for others.

Theories of Between-Issue Variation

Finally, despite an abundance of theories for why party elite influence might vary across policy issues, there is a distinct lack of evidence as to the empirical magnitude of this variation. I have provided a quantitative estimate of this variation and placed it in the context of existing between-study variation, thus revealing there is substantively large variation across policy issues. This underlines the importance of testing and refining theories to explain variation across policy issues in party elite influence, since there is a great deal of variation to explain. My results offer evidence consistent with one such theory: that "pretreatment" by the party cue explains variation across policy issues in party elite cue effects (Gaines, Kuklinski, and Quirk 2007; Slothuus 2016). Furthermore, I uncovered some evidence of differences in party elite cue effects across broad policy domains, consistent with theories that focus on the different character of policy issues (e.g., Carmines and Stimson 1980). This provides new evidence in line with other recent empirical work (e.g., Chong and Mullinix 2019; Clifford, Leeper, and Rainey 2021; but see Hill and Huber 2019).

Limitations and Closing Thoughts

Before closing, I note some limitations of my results. My estimates of the between-issue range in party elite cue effects are themselves somewhat variable across analyses, ranging from 6 percent to 27 percent of the opinion scale. On the one hand, this is surely partly due to sampling variability: the uncertainty intervals on my estimates of the range are typically large, owing to the empirical demands of trying to estimate a distribution of average treatment effects. Thus, while four out of five of the estimates that I report fall between 14 percent and 27 percent of the outcome scale, this should not be taken as the final word on the range in effect sizes that is due to policy issues alone.

On the other hand, the variability in my estimates could also suggest that heterogeneity in party elite cue effects across policy issues depends systematically on other factors that vary across my analyses—such as the type of cue (e.g., party leader versus generic elite) and the sample of respondents (e.g., more versus less attentive). This raises intriguing research questions. For example, are certain types of people more discerning than others in the policy issues on which they choose to follow party elite cues (i.e., higher between-issue variation)? A large body of existing research focuses on differences between people in the *average* effect of party elite cues (see Bullock 2020 for a review); the politically knowledgeable, for instance. By expanding the samples of policy issues that they study, scholars could also ask whether such individuals are more *discerning* in their cue following. I leave these and similar questions for future work.

We have learned much about party elite influence on public opinion, but there remains a great deal still to learn. The results reported in this paper suggest that this learning will be more productive and efficient if we significantly expand the samples of policy issues included in our studies.

Supplementary Material

Supplementary Material may be found in the online version of this article: https://doi.org/10.1093/poq/nfac052.

Data Availability

Replication data and documentation are available at https://osf.io/t2bpj/.

References

Adams, James, Michael Clark, Lawrence Ezrow, and Garrett Glasgow. 2004. "Understanding Change and Stability in Party Ideologies: Do Parties Respond to Public Opinion or to Past Election Results?" *British Journal of Political Science* 34:589–610.

Agadjanian, Alexander. 2020. "When Do Partisans Stop Following the Leader?" *Political Communication* 38:351–69.

- Bakker, Bert N., Yphtach Lelkes, and Ariel Malka. 2020. "Understanding Partisan Cue Receptivity: Tests of Predictions from the Bounded Rationality and Expressive Utility Perspectives." *Journal of Politics* 82:1061–77.
- Barber, Michael, and Jeremy C. Pope. 2019. "Does Party Trump Ideology? Disentangling Party and Ideology in America." *American Political Science Review* 113:38–54.
- Boudreau, Cheryl, and Scott A. MacKenzie. 2014. "Informing the Electorate? How Party Cues and Policy Information Affect Public Opinion about Initiatives." *American Journal of Political Science* 58:48–62.
- Brader, Ted A., and Joshua A. Tucker. 2012. "Following the Party's Lead: Party Cues, Policy Opinion, and the Power of Partisanship in Three Multiparty Systems." *Comparative Politics* 44:403–15.
- Broockman, David E., and Daniel M. Butler. 2017. "The Causal Effects of Elite Position-Taking on Voter Attitudes: Field Experiments with Elite Communication." *American Journal of Political Science* 61:208–21.
- Bullock, John G. 2011. "Elite Influence on Public Opinion in an Informed Electorate." American Political Science Review 105:496–515.
- 2020. "Party Cues." In *The Oxford Handbook of Electoral Persuasion*, edited by Elizabeth Suhay, Bernard Grofman, and Alexander H. Treschel, 129–150. New York, NY, USA: Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190860806.001.0001/oxfordhb-9780190860806-e-2.
- Bürkner, Paul-Christian. 2017. "Brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80:1–28.
- Carmines, Edward G., and James A. Stimson. 1980. "The Two Faces of Issue Voting." American Political Science Review 74:78–91.
- Chong, Dennis, and Kevin J. Mullinix. 2019. "Information and Issue Constraints on Party Cues." American Politics Research 47:1209–38.
- Ciuk, David J., and Berwood A. Yost. 2016. "The Effects of Issue Salience, Elite Influence, and Policy Content on Public Opinion." *Political Communication* 33:328–45.
- Clifford, Scott, Thomas J. Leeper, and, Carlisle Rainey. 2021. "Generalizing Survey Experiments Using Topic Sampling: An Application to Party Cues". *Paper presented at the American Political Science Association's Annual Meeting 2019*, Washington DC. http://scottaclifford. com/wp-content/uploads/2012/02/CliffordLeeperRainey_Topic-Sampling.pdf.
- Clifford, Scott, Geoffrey Sheagley, and, Spencer Piston. 2021. "Increasing Precision without Altering Treatment Effects: Repeated Measures Designs in Survey Experiments." *American Political Science Review* 115:1048–65.
- Cohen, Geoffrey L. 2003. "Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs." *Journal of Personality and Social Psychology* 85:808–22.
- Coppock, Alexander. 2019. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* 7:613–28.
- Coppock, Alexander, Seth J. Hill, and Lynn Vavreck. 2020. "The Small Effects of Political Advertising Are Small Regardless of Context, Message, Sender, or Receiver: Evidence from 59 Real-Time Randomized Experiments." *Science Advances* 6:eabc4046.
- Coppock, Alexander, and Oliver McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research & Politics* 6:205316801882217.
- Ehret, Phillip J., Leaf Van Boven, and David K. Sherman. 2018. "Partisan Barriers to Bipartisanship: Understanding Climate Policy Polarization." Social Psychological and Personality Science 9:308–18.
- Findley, Michael G., Kyosuke Kikuta, and Michael Denly. 2021. "External Validity." Annual Review of Political Science 24:365–93.

- Gaines, Brian J., James H. Kuklinski, and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15:1–20.
- Gelman, Andrew, and Jennifer Hill. 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge, UK: Cambridge University Press.
- Guisinger, Alexandra, and Elizabeth N. Saunders. 2017. "Mapping the Boundaries of Elite Cues: How Elites Shape Mass Opinion across International Issues." *International Studies Quarterly* 61:425–41.
- Hill, Seth J., and Gregory A. Huber. 2019. "On the Meaning of Survey Reports of Roll-Call 'Votes." American Journal of Political Science 63:611–25.
- Kam, Cindy D. 2005. "Who Toes the Party Line? Cues, Values, and Individual Differences." *Political Behavior* 27:163–82.
- Krosnick, Jon A. 1990. "Government Policy and Citizen Passion: A Study of Issue Publics in Contemporary America." *Political Behavior* 12:59–92.
- Lenz, Gabriel S. 2012. Follow the Leader? How Voters Respond to Politicians' Policies and Performance. Chicago, IL, USA: University of Chicago Press.
- Linden, Audrey Helen, and Johannes Hönekopp. 2021. "Heterogeneity of Research Results: A New Perspective From Which to Assess and Promote Progress in Psychological Science." *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 16:358–76.
- McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. Boca Raton, FL, USA: CRC Press.
- Mooney, Christopher Z. 2001. The Public Clash of Private Values: The Politics of Morality Policy. London, UK: SAGE Publications.
- Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2:109–38.
- Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." American Political Science Review 113:517–29.

Nicholson, Stephen P. 2012. "Polarizing Cues." American Journal of Political Science 56:52-66.

Petersen, Michael Bang, Martin Skov, Søren Serritzlew, and Thomas Ramsøy. 2013. "Motivated Reasoning and Political Parties: Evidence for Increased Processing in the Face of Party Cues." *Political Behavior* 35:831–54.

- Pollock, Philip H., Stuart A. Lilie, and M. Elliot Vittes. 1993. "Hard Issues, Core Values and Vertical Constraint: The Case of Nuclear Power." *British Journal of Political Science* 23:29–50.
- Sevenans, Julie. 2021. "How Public Opinion Information Changes Politicians' Opinions and Behavior." *Political Behavior* 43:1801–23.
- Slothuus, Rune. 2016. "Assessing the Influence of Political Parties on Public Opinion: The Challenge from Pretreatment Effects." *Political Communication* 33:302–27.
- Tappin, Ben M., and Luke B. Hewitt. 2021. "Estimating the Persistence of Party Cue Influence in a Panel Survey Experiment." *Journal of Experimental Political Science* 1–12. https://doi. org/10.1017/XPS.2021.22.
- Tappin, Ben M., and Ryan McKay. 2021. "Estimating the Causal Effects of Cognitive Effort and Policy Information on Party Cue Influence." PsyArxiv 1–44. https://psyarxiv.com/tdk3y/.
- Ternovski, J., and L. Orr. 2022. "A Note on Increases in Inattentive Online Survey-Takers Since 2020." Journal of Quantitative Description: Digital Media 2. https://doi.org/10.51685/jqd.2022.002.
- Van Boven, Leaf, Phillip J. Ehret, and David K. Sherman. 2018. "Psychological Barriers to Bipartisan Public Support for Climate Policy." *Perspectives on Psychological Science : A Journal of the Association for Psychological Science* 13:492–507.
- Yarkoni, Tal. 2020. "The Generalizability Crisis." Behavioral and Brain Sciences 45:E1. https:// doi.org/10.1017/S0140525X20001685.
- Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge, UK: Cambridge University Press.