Preferential Abstention in Conjoint Experiments

Abstract

Conjoint experiments are often used to mimic political choices that people face, such as voting for public officials or selecting news stories. Conjoint designs, however, do not always mirror the real-world decision-making contexts that individuals engage in because respondents are typically forced to select one of the available options. Theoretically, we illustrate how offering respondents an abstention option can produce average marginal component effects (AMCEs) of differing signs and magnitudes relative to a forced-choice outcome. This difference depends on 1) the proportion of respondents who would rather abstain than select profiles lacking their preferred attribute-levels, and 2) those respondents' preference orderings. Empirically, we replicate two conjoint experiments and demonstrate how omitting a realistic abstention option could lead to different AMCE estimates.

1 Introduction

Social scientists frequently use conjoint experiments to study a wide array of decisions that humans encounter. Conjoint experiments ask respondents to evaluate two or more alternatives that differ along a set of characteristics, often requiring participants to indicate which of the available alternatives they most prefer.[1] This "forced-choice" design has been used to measure individuals' preferences over candidates in elections (Carlson 2015), economic policies (Chilton, Milner and Tingley 2020), and news stories (Mukerjee and Yang 2020). Yet, participants may ordinarily abstain from these types of decision-making processes entirely in daily life if they find none of the choices suitable.

We outline theoretically how the average marginal component effects (AMCEs) associated with attribute-levels in a forced-choice design can differ in sign *or* magnitude relative to the AMCEs for the same attribute-levels in a design allowing for abstention. Importantly, researchers cannot know beforehand how any of the AMCEs obtained in a forced-choice design compare to the estimates that would arise from a design with an abstention option. We then replicate two conjoint experiments that used forced-choice outcomes (Funck and McCabe 2022; Mummolo 2016) to investigate how including an abstention option could yield statistically and substantively different conclusions.

2 Forced-Choice Outcomes, Abstention, and AMCEs

The standard conjoint experiment asks each respondent ($i \in 1, ..., N$) to evaluate a fixed number of tasks ($k \in 1, ..., K$). Each task presents respondents with a fixed number of profiles ($j \in 1, ..., J$) that consist of randomly assigned levels for each attribute ($l \in 1, ..., L$).

After viewing the profiles in each task, respondents express their preferences toward those profiles, most commonly with a "forced-choice" outcome that requires respondents to indicate the profile they most prefer. These choices are then typically used to calculate AMCEs, which represent the probability a profile with a given attribute-level is selected relative to a randomly-generated profile with that attribute's baseline level (Hainmueller, Hopkins and Yamamoto 2014, 19).[2]

A key assumption of the conjoint framework is that "respondents must choose one preferred profile j within each choice task k" (Hainmueller, Hopkins and Yamamoto, 2014, 7). However, for many of the decision-making contexts to which conjoint experiments are applied, ample evidence exists that many people abstain when prompted to make a choice because they do not hold an opinion or do not approve of their choices. For instance, individuals' decision to cast a ballot for one of the available candidates or parties in an election is far from universal, even in countries with compulsory voting (Blais 2006).

Notably, abstention is often non-random and certain *types* of respondents, as defined by their preference orderings, might be more likely to abstain. For instance, partisans' willingness to consume news is influenced by whether their party is advantaged or disadvantaged by salient events (Kim and Kim 2021). Specific to survey taking, when given the option to respond "don't know", individuals in lower socioeconomic strata, women, and racial minorities may be more likely to abstain (Berinsky 2008).[3] As such, when forced-choice outcomes are employed in conjoint experiments, it is reasonable to anticipate that some respondents may be artificially induced to make a choice. We focus on the mechanical issues that arise when we force respondents' to elicit preferences and how it may yield different estimates of the AMCEs than would manifest if respondents could abstain.[4]

2.2 Implications of a Forced-Choice Outcome on AMCEs

We adapt a stylized conjoint experiment from Abramson, Koçak and Magazinnik (2022) in which $N = 5$ voters are asked to evaluate pairs of candidates based on their positions on two policy issues, which serve as our *attributes*. For each issue, candidates can take one of two positions, which constitute our *levels*. First, say on abortion policy, voters can be pro-life (PL) or pro-choice (PC). Second, on tax policy, voters can support increasing (I) or decreasing (D) taxes on the upper-class. For simplicity, we assume all voters prioritize candidates' abortion policy stances to candidates' tax policy stances, and that voters' abortion and tax policy preferences are ideologically coherent. In other words, all voters preferring pro-life candidates ($PL > PC$) are conservatives and therefore also prefer candidates pledging to cut taxes ($D > I$). To start, let three of the voters be conservatives and two of the voters be liberals. The full preference ordering for the four possible candidate types for the five voters is presented in Table 1a.

We use the voter preference rank-orderings in Table 1a to determine the electorate's aggregate preferences for each unique combination of candidate profiles in Table 1b. The final four columns of the table provide the vote tallies for each candidate comparison under different forced-choice scenarios: (1) when voters must vote for the candidate they prefer ("No Abstentions"); (2) when all voters abstain if neither candidate possesses their most preferred level of their prioritized attribute, abortion policy ("Uniform Abstentions"); (3) when voters 1, 3, and 5 abstain if neither candidate is pro-life, but voters 2 and 4 always vote ("Pro-life Abstentions"); and (4) when voters 2 and 4 abstain if neither candidate is pro-choice, but voters 1, 3, and 5 always vote ("Pro-choice Abstentions"). In the first scenario, the first candidate profile wins each comparison by a 3 to 2 margin, but if abstentions are allowed, the winning candidate and margin of victory differ only when both candidates share the same abortion policy position (Table 1b, Rows 3 and 6).

Finally, as in Abramson, Koçak and Magazinnik (2022), we use Proposition 3 in Hainmueller, Hopkins and Yamamoto (2014, 16) to calculate the AMCEs. To do so, we first obtain the difference in the number of votes a candidate with one level of an attribute

Table 1: Individual and aggregate preferences over candidate profiles with differing policy positions.

(a) Voter preferences

Rank	V1	V2	V3	V4	V5
1	PL, D	PC, I	PL, D	PC, I	PL, D
2	PL, I	PC, D	PL, I	PC, D	PL, I
3	PC, D	PL, I	PC, D	PL, I	PC, D
4	PC, I	PL, D	PC, I	PL, D	PC, I

(b) Aggregate preferences

	Comparison	V1	V2	V3	V4	V5	No Abstentions	Uniform Abstentions	Pro-Life Abstentions	Pro-Choice Abstentions
1	PL, D vs. PC, I	PL, D	PC, I	PL, D	PC, I	PL, D	3,2	3,2	3,2	3,2
2	PL, D vs. PC, D	PL, D	PC, D	PL, D	PC, D	PL, D	3,2	3,2	3,2	3,2
3	PL, D vs. PL, I	PL, D	PL, I	PL, D	PL, I	PL, D	3,2	3,0	3,2	3,0
4	PL, I vs. PC, I	PL, I	PC, I	PL, I	PC, I	PL, I	3,2	3,2	3,2	3,2
5	PL, I vs. PC, D	PL, I	PC, D	PL, I	PC, D	PL, I	3,2	3,2	3,2	3,2
6	PC, D vs. PC, I	PC, D	PC, I	PC, D	PC, I	PC, D	3,2	0,2	0,2	3,2

Notes: PL=Pro-life, PC=Pro-choice, D=Decrease upper-class taxes, I=Increase upper-class taxes. In the four rightmost columns of Table 1b, the first and second numbers indicate the number of voters preferring the first and second candidates in the comparison pair given respondents' ability to abstain (as indicated by the column headings).

would receive compared to a candidate with the other level of that same attribute, holding the second attribute constant, when pitted against each possible candidate. Following Hainmueller, Hopkins and Yamamoto (2014, 7), we code votes as 1 when a candidate is selected and 0 when it is not. We then sum these differences and normalize the sums by the product of the number of possible profiles (4), number of possible profiles with a fixed level of one of the two attributes (2) and the number of voters (5). So, the denominator is calculated as the number of unique profiles times the number of voters times the number of possible profiles with the unique levels of copartisanship and corruption (i.e., $4 \times 5 \times 2$).

The AMCEs for each scenario are shown in Table 2. When using a forced-choice outcome that requires all voters to cast a ballot, the AMCE for pro-life is 0.10 (Table 2a, Column 1) and the AMCE for cutting taxes is 0.05 (Table 2b, Column 1). However, when voters are allowed to abstain if neither candidate in the matchup has their preferred abortion policy stance, the AMCE for pro-life increases to 0.15, while the AMCE for cutting taxes decreases to 0.025 (Table 2, Column 2).

To understand why the AMCEs change in opposite directions under uniform abstention, we focus on the matchups in which the vote tallies diverge under the forced-choice and uniform abstention scenarios. For instance, for the pro-life AMCE, consider the comparison of $\bar{Y}(PL, D; PC, I)$ - $\bar{Y}(PC, D; PC, I)$ (Table 2a, Row 4); voters' aggregate preference for pro-life candidates is stronger when abstention is allowed because pro-life voters no longer cast votes for pro-choice candidates when no pro-life candidates are present (e.g., the $PC, D; PC, I$ matchup). Put differently, voters no longer dilute their preferences over candidates' abortion policy stances by selecting candidates who do not share their own position.

Table 2: AMCEs for abortion and tax policies varying abstention options.

(a) Pro-life

		No Abstention	Uniform Abstention	Pro-life Abstentions	Pro-choice Abstentions
1	$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PC,D;PL,D)$	0.50	-0.50	0.50	-0.50
2	$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PC,D;PL,I)$	1	1	1	1
3	$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PC,D;PC,D)$	0.50	2	2	0.50
4	$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PC,D;PC,I)$	0	3	3	0
5	$\bar{Y}(PL,I;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0	-2	0	-2
6	$\bar{Y}(PL,I;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0.50	-0.50	0.50	-0.50
7	$\bar{Y}(PL,I;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	1	1	1	1
8	$\bar{Y}(PL,I:PC,I)$ - $\bar{Y}(PC,I;PC,I)$	0.50	2	2	0.50
	AMCE	$=\frac{4}{40}=0.10$	$=\frac{6}{40}=0.15$	$=\frac{10}{40}=0.25$	$=\frac{0}{40}=0.00$

(b) Cutting taxes

		No Abstention	Uniform Abstention	Pro-life Abstentions	Pro-choice Abstentions
1	$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PL,I;PL,D)$	0.50	1.50	0.50	1.50
2	$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PL,I;PL,I)$	0.50	1.50	0.50	1.50
3	$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PL,I;PC,D)$	0	0	0	0
4	$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PL,I;PC,I)$	0	0	0	0
5	$\bar{Y}(PC,D;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0	0	0	0
6	$\bar{Y}(PC,D;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0	0	0	0
7	$\bar{Y}(PC,D;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	0.50	-1	-1	0.50
8	$\bar{Y}(PC,D;PC,I)$ - $\bar{Y}(PC,I;PC,I)$	0.50	-1	-1	0.50
	AMCE	$=\frac{2}{40}=0.05$	$=\frac{1}{40}=0.025$	$=\frac{-1}{40}=-0.025$	$=\frac{4}{40}=0.10$

Notes: PL=Pro-life, PC=Pro-choice, D=Decrease upper-class taxes, I=Increase upper-class taxes.

Alternatively, for the cutting taxes AMCE, because decreasing taxes is linked with being pro-choice, when there is uniform abstention along the abortion dimension there are fewer voters that prefer to cut taxes contributing to the AMCE (Table 2a, Rows 7 and 8). For instance, in the $\bar{Y}(PC,D;PC,I)$ - $\bar{Y}(PC,I;PC,I)$ comparison, voters who prefer a pro-life candidate will not vote if there is a pro-choice candidate, which reduces the aggregate preference toward cutting taxes.

In the final two columns of Table 2, we consider how the pro-life and cutting taxes AMCEs change when only pro-life or only pro-choice voters abstain. First, when pro-choice voters always vote but pro-life voters abstain if no candidates espouse pro-life positions, we see that the pro-life AMCE increases further to 0.25 (Table 2a, Column 3). Pro-choice voters diminish their aggregate preference for pro-choice candidates by casting votes for pro-life candidates when no pro-choice candidates are available. Yet, pro-life voters do not vote when confronted with only pro-choice candidates.

The cutting taxes AMCE, however, decreases to -0.025 indicating that the electorate's aggregate preference on tax policy is in support of candidates pledging to raise taxes on the upper-class *despite* a majority of voters preferring candidates pledging to cut those taxes. Pro-life voters no longer express their preference for pro-choice candidates pledging to cut taxes (Table 2b, Column 3, Rows 7 and 8), but pro-choice voters continue to cast ballots for pro-life candidates pledging to raise taxes. In aggregate, these expressed preferences yield a negative, rather than positive, AMCE for cutting taxes.

Finally, when only pro-choice voters abstain, we observe that the AMCE for pro-life is 0.00 even though the majority of voters prefer pro-life candidates. This quantity arises because pro-choice voters no longer vote when only pro-life candidates are present (Table 2a, Column 4, Rows 1, 5, and 6), but pro-life voters still cast ballots for pro-choice candidates when no pro-life candidates are available. Meanwhile, the AMCE for cutting taxes increases to 0.10 as pro-choice voters no longer express their preference for increasing taxes by voting for pro-life candidates pledging to raise taxes (Table 2b, Column 4, Rows 1 and 2). As pro-life voters outnumber pro-choice voters, these abstentions increasingly boost the pro-life AMCE. These examples of uniform and asymmetric abstention showcase that the consequences of using forced-choice outcomes in conjoint experiments are not straightforward. Both the *magnitudes* and *signs* of AMCEs obtained using a forced-choice outcome may substantively differ from a design in which respondents are not forced to express preferences they do not hold.

To illustrate how the rate of abstention among voters with different preference orderings impacts the AMCEs, we extend our example to N-voters in the Supplementary Material. We find that the differences in the AMCEs are based on the distribution of persons in the sample with each preference ordering and the rate of abstention among those who abstain if none of the profiles presented contain the attribute-level they prioritize. In other words, these differences depend on *which type* of respondents would abstain, as described by their preference orderings, and *how many* respondents of each type would abstain.[5]

Unfortunately, researchers cannot know beforehand which types of respondents, as defined by their preference orderings, would rather abstain than select profiles lacking their preferred attribute-levels. Further, if respondents with certain preference orderings are more likely to abstain, forced-choice outcomes compel respondents to reveal preferences they would otherwise hold. Therefore, researchers cannot rely on any “rules of thumb" to speculate ex post on whether the directionality or magnitude of their AMCEs from forced-choice responses generalize to realistic contexts when abstention is possible.

3 Replication & Extension of Forced-Choice Conjoint

We explore empirically how forced-choice outcomes could produce different estimates of the AMCEs relative to a choice set allowing for abstention by replicating two published conjoint experiments. The first study examines how the complexity of the information environment affects the impact of scandals on vote choice (Funck and McCabe 2022, henceforth "F-M"). The second study, described in the Supplementary Materials, explores how the topic relevance and source partisanship of news stories impact individuals' media consumption (Mummolo 2016).

In their study, F-M ask American respondents recruited through Lucid to view pairs of hypothetical candidates running for U.S. Congress. Partisanship is fixed at the task-level to mirror a general election, with one candidate randomly presented as a Democrat and the other as a Republican. For each task, respondents indicate the candidate for whom they would prefer to vote using a forced-choice outcome.

The two experimental attributes of interest for F-M's primary "Information Hypothesis" are (1) whether candidates are accused of improper behavior and (2) the amount of information provided about the candidates. First, F-M randomly assign one of six levels of "Recent news" for each candidate; three levels are neutral or positive ("No recent news", "Recently honored for public service", or "Recently celebrated wedding anniversary"), while the others implicate the candidate in a scandal ("Recently accused of sexual harassment", "Recently accused of cheating on spouse", or "Recently accused of leaking confidential information").

Second, the number of attributes provided for both candidates is randomly manipulated at the task-level. In the "Low Information" condition, respondents receive only the candidates' party affiliation and recent news. In the "Medium" and "High Information" conditions, respondents receive randomly assigned levels of three or eight additional attributes of the candidates. In line with their "Information Hypothesis", F-M find voters are less likely to select candidates associated with scandalous allegations, but the magnitude of this penalty shrinks as the information environment's complexity increases.

We replicate F-M's experiment using 2,254 respondents from Amazon's Mechanical Turk (MTurk). We match F-M's protocol with a few modifications. First, half of our respondents were assigned to a *forced-choice arm*. These participants were required to indicate which candidate they preferred in each task. The other half of our respondents were placed in a *abstention option arm* where they could indicate their preferred candidate or abstain from making a selection. Second, after completing their 6 tasks, respondents were presented again with the first profile pair they evaluated but were provided the outcome measure from the opposite experimental arm (e.g., forced-choice arm respondents were given the option to abstain).[6]

Following F-M, we estimate our AMCEs using ordinary least squares regression with standard errors clustered by respondent. Our outcome is a binary indicator for whether a respondent selected an available profile, coded as "1", and all other profiles in the task are coded as "0." If a respondent abstained, all profiles in the task are coded as "0." Importantly, respondents abstain often when allowed to do so. In our abstention arm, respondents selected none of the profiles in 25.8% (1764) of tasks in that arm and 31.0% (699) of respondents abstained at least once.[7] This descriptive finding highlights that response options in choice-based conjoint experiments may induce meaningfully different behavior from respondents. We interact all non-baseline attribute-levels with a binary indicator for each respondent's outcome measure arm so we can compare directly the AMCEs yielded when respondents are forced to choose a candidate versus when they are allowed to abstain. Finally, because F-M's hypotheses implicate differences in the AMCEs recovered for attribute-levels of "Recent news" across facets of information environment complexity, we focus on those differences as opposed to the underlying AMCEs.[8]

3.2 Forced-Choice Design Can Result in Different Conclusions

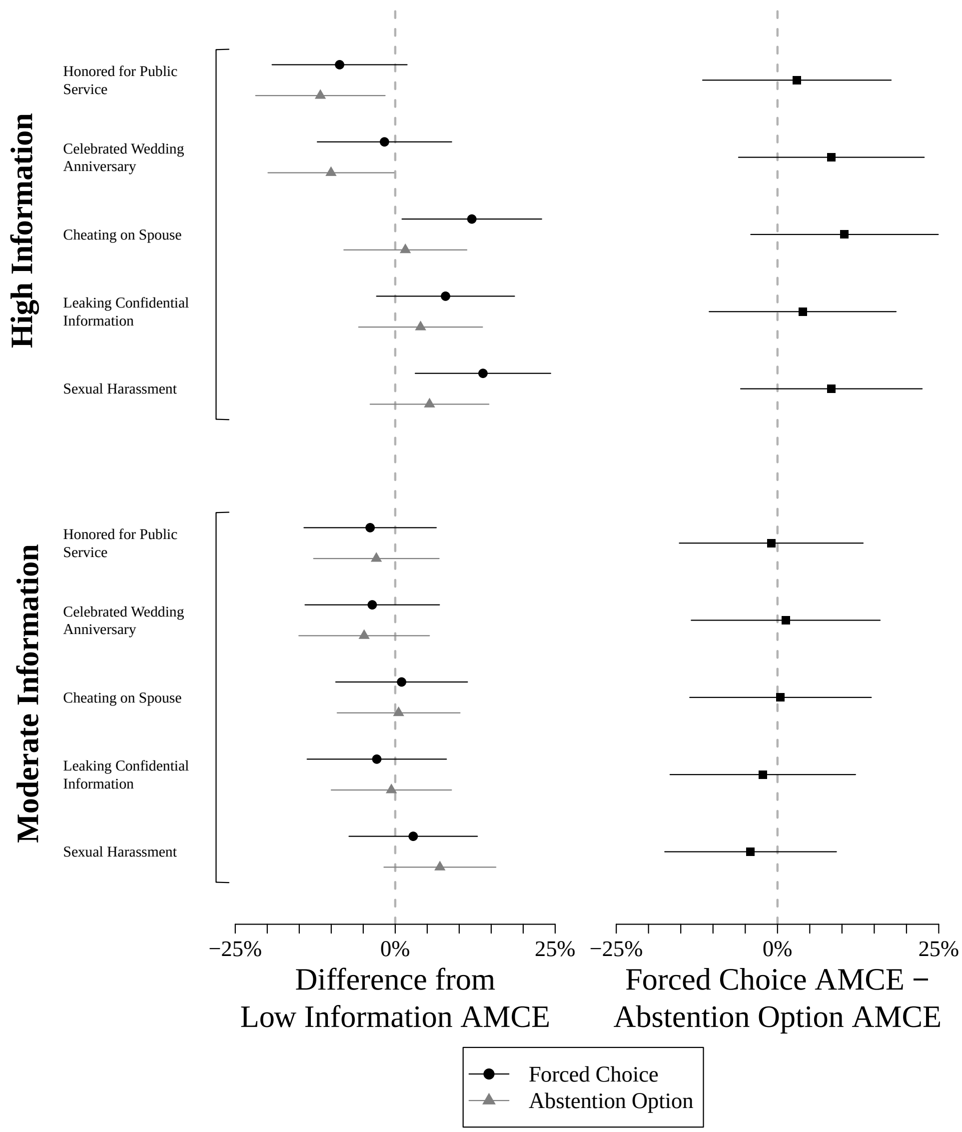
In Figure 1, the left pane displays the difference between the AMCE for each non-baseline level of the "Recent news" attribute completed in the "Low Information" environment and the corresponding attribute-level in the "Moderate" and "High Information" environments among respondents in the forced-choice (black circles) and abstention option (grey triangles) arms. The right pane of Figure 1 presents each of the differences in AMCEs between the estimates yielded by the forced-choice and abstention arms (black squares). First, when comparing the "Low" and "Moderate Information" conditions, we see that all of the differences corresponding with the forced-choice and abstention-option arms lead to substantively similar conclusions. In other words, we do not replicate F-M's initial results; we find no difference in how much respondents prefer any "Recent News" attribute-level between the two lowest information categories.

Some distinctions emerge when comparing the "Low" and "High Information" conditions in each of the two arms. For two of the three negative events ("Cheating on spouse" and "Sexual harassment"), our forced-choice arm replicates F-M's finding—respondents penalize candidates less for transgressions in the "High Information" condition than in the "Low Information" condition (i.e. respondents are more likely to vote for corrupt candidates in the "High" compared to the "Low Information" condition)—but these differences do not manifest in our abstention option arm. However, we cannot reject the null hypothesis that the estimates associated with the forced-choice and abstention arms are not different from each other. The right pane of Figure 1 shows that the Bonferroni-corrected 95% confidence intervals for the differences in the differences of the AMCEs across the two arms all include zero.[9]

Similarly, while the two positive news stories ("Honored for public service" and "Celebrated wedding anniversary") are not of theoretical relevance for F-M's "Information Hypothesis," the differences in AMCEs associated with these attribute-levels between the "Low" and "High Information" conditions are not statistically distinguishable in the forced-choice arm, but are distinguishable in the abstention arm (though, again, the differences in these estimates across arms are not themselves distinguishable).

In the Supplementary Materials, we consider one potential explanation for why the differences in AMCEs for negative "Recent News" attribute-levels across information conditions are smaller (albeit not statistically distinction) in our abstention arm relative to the forced-choice arm: respondents who are forced to choose are reluctant to select for scandalized candidates in low information environments, but are more willing to do so when additional information can obscure or compensate for scandals. Yet, when respondents can abstain and prioritize candidates' reputations, they decline to make a selection when presented with scandalized candidates irrespective of other information. In the Supplementary Materials, we probe the profile- and task-level characteristics associated with abstention and find that respondents are more likely to abstain when both candidates or the candidate who shares their party affiliation are implicated in scandals.

Figure 1: Effect of scandalous news and information complexity on vote choice.



Notes: The left pane of this figure presents the differences in the AMCE for each non-baseline "Recent News" attribute-level for conjoint tasks situated in a "Low Information" conjoint task relative to when the same attribute-level is situated in "Moderate Information" or "High Information." Black circles and grey triangles represent the differences for each attribute-level among respondents in the forced-choice and abstention arms, respectively. The right pane of the figure displays for each attribute-level in each information environment the difference in the differences in AMCEs obtained in the forced-choice and abstention arms. Lines in both panes represent Bonferroni-adjusted 95% confidence intervals ($\alpha = \frac{0.05}{20} = 0.0025$) for each attribute-level among respondents in the forced-choice arm (Liu and Shiraito 2023). See the Supplemental Materials for full model summary.

Further, we find that as the information environment becomes more complex, abstention becomes less common only if the respondent's copartisan is scandalized; otherwise, the probability of abstention increases as the information environment becomes more complex.

4 Conclusion

We have shown theoretically and empirically that forcing participants to reveal a preference when they would otherwise prefer to abstain may lead to AMCEs of different magnitudes or signs than would be recovered if an abstention option were offered. In some contexts, such as when political elites must hire civil servants (Liu 2019) or select which lawmakers to lobby (Miller 2022), abstention is unlikely and undesirable because a choice *must* be made as a key function of the respondents' occupations. Additionally, when researchers are chiefly interested in measuring respondents' underlying preferences as opposed to how they express those preferences in decision-making contexts and they have reason to believe respondents hold preferences on their concept of interest, using rating outcomes or omitting abstention options may be appropriate (e.g., Ganter, 2023). However, we advise that researchers interested in modeling respondents' expressed preferences in decision-making contexts provide an abstention option if the data generating process naturally includes one. For example, some studies have included an abstention option to model the selection over candidates in elections (Agerberg 2020) and locations to migrate (Ghosn et al. 2021). Including an abstention option is particularly important in light of the non-negligible abstention rates observed among the respondents in our abstention arms—across both replications, respondents abstained in approximately one-fourth of all tasks and roughly one-third of respondents abstained at least once.

If researchers wish to use forced-choice designs, we encourage them to be explicit about the contexts to which their findings might apply. The AMCEs recovered with a forced-choice outcome reflect the revealed preferences of respondents *when a choice is required*. Alternatively, if researchers are concerned with modeling the entire decision-making process, it may be better to modify the research design to make use of an appropriate estimating procedure that accounts for task attributes that impact respondents' decision to select as well as preferences if a selection is made.

Notes

1. We identify and hand-code the outcome measures of all conjoint experiments published in the *American Journal of Political Science*, *American Political Science Review*, and *Journal of Politics* as of June 2022 in the Supplementary Materials. Of the published conjoint studies, 88% include a forced-choice outcome. The second most common outcome measure in conjoint designs is the individual rating, whereby respondents evaluate each profile using an ordinal scale (Bansak et al. 2021). The issues we describe concerning forced-choice outcomes do not apply to rating outcomes because respondents rate individual profiles rather than choose among profiles.

2. While political scientists sometimes estimate other quantities of interest, such as marginal means (Leeper, Hobolt and Tilley 2020), AMCEs are the most common estimates presented in conjoint studies.

3. Including abstention options (e.g., "don't know") in outcome measures is sometimes discouraged because respondents may use these options to satisfice and conceal real opinions (Krosnick et al. 2002). The survey questions that typically fall prey to this concern strive to measure latent concepts that most ordinary people have preferences over, such as approval of politicians and policies.

4. The disconnect between preference options and preference expression also sparks concerns about the external validity of the forced-choice design. Within the validity typology proposed by Egami and Hartman (2022), these conceptual issues fall under Y-validity, or how well the outcome measures map onto the theoretical quantities of interest.

5. We also discuss in the Supplementary Materials how varying other facets of our conjoint example can affect the resulting AMCEs—namely, whether different voter types have similar or different preferences between the levels of each attribute, and the degree to which they similarly prioritize attributes when rank-ordering potential candidate profiles.

6. We discuss how respondents' choices in the final task correspond to their choices when they evaluated the same task with the opposite outcome question in the Supplementary Materials. Among respondents in the forced-choice arm that completed the entire survey, approximately one third abstained when allowed to do so in the final task (27.6% F-M, 33.9% Mummolo).

7. Similarly, in our replication of Mummolo, 2016, respondents abstained in 24.3% (2335) of tasks and 31.1% (565) of respondents abstained at least once.

8. The Supplementary Materials include a presentation of the underlying AMCEs.

9. The differences in the differences of the AMCEs across the forced-choice and abstention arms for the High Information “Cheating on spouse” and “Sexual harassment” attribute-levels are statistically distinguishable at $p \leq 0.05$ and $p \leq 0.10$, respectively, before Bonferroni corrections are implemented, but do not retain distinguishability once they are applied.

References

Abramson, Scott F, Korhan Koçak and Asya Magazinnik. 2022. "What Do We Learn About Voter Preferences from Conjoint Experiments." *American Journal of Political Science* .

Agerberg, Mattias. 2020. "The Lesser Evil? Corruption Voting and the Importance of Clean Alternatives." *Comparative Political Studies* 53(2):253–287.

Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins and Teppei Yamamoto. 2021. Conjoint Survey Experiments. In *Advances in Experimental Political Science.* Cambridge University Press.

Berinsky, Adam J. 2008. Survey Non-Response. In *The SAGE Handbook of Public Opinion Research*, ed. Wolfgang Donsbach and Michael W. Traugott. Sage pp. 309–321.

Blais, André. 2006. "What Affects Voter Turnout?" *Annual Review of Political Science* 9:111–125.

Carlson, Elizabeth. 2015. "Ethnic Voting and Accountability in Africa: A Choice Experiment in Uganda." *World Politics* 67(2):353–385.

Chilton, Adam S, Helen V Milner and Dustin Tingley. 2020. "Reciprocity and Public Opposition to Foreign Direct Investment." *British Journal of Political Science* 50(1):129–153.

Egami, Naoki and Erin Hartman. 2022. "Elements of External Validity: Framework, Design, and Analysis." *American Political Science Review* pp. 1–19.

Funck, Amy S and Katherine T McCabe. 2022. "Partisanship, Information, and the Conditional Effects of Scandal on Voting Decisions." *Political Behavior* 44:1389–1409.

Ganter, Flavien. 2023. "Identification of preferences in forced-choice conjoint experiments: Reassessing the quantity of interest." *Political Analysis* 31(1):98–112.

Ghosn, Faten, Tiffany S Chu, Miranda Simon, Alex Braithwaite, Michael Frith and Joanna Jandali. 2021. "The Journey Home: Violence, Anchoring, and Refugee Decisions to Return." *American Political Science Review* 115(3):982–998.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Kim, Jin Woo and Eunji Kim. 2021. "Temporal Selective Exposure: How Partisans Choose When to Follow Politics." *Political Behavior* 43(4):1663–1683.

Krosnick, Jon A, Allyson L Holbrook, Matthew K Berent, Richard T Carson, W Michael Hanemann, Raymond J Kopp, Robert Cameron Mitchell, Stanley Presser, Paul A Ruud and V Kerry Smith. 2002. "The Impact of 'No Opinion' Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66(3):371–403.

Leeper, Thomas J, Sara B Hobolt and James Tilley. 2020. "Measuring subgroup preferences in conjoint experiments." *Political Analysis* 28(2):207–221.

Liu, Guoer and Yuki Shiraito. 2023. "Multiple Hypothesis Testing in Conjoint Analysis." *Political Analysis* .

Liu, Hanzhang. 2019. "The Logic of Authoritarian Political Selection: Evidence from a Conjoint Experiment in China." *Political Science Research and Methods* 7(4):853–870.

Miller, David R. 2022. "On Whose Door to Knock? Organized Interests' Strategic Pursuit of Access to Members of Congress." *Legislative Studies Quarterly* 47(1):157–192.

Mukerjee, Subhayan and Tian Yang. 2020. "Choosing to Avoid? A Conjoint Experimental Study to Understand Selective Exposure and Avoidance on Social Media." *Political Communication* 38(3):222–240.

Mummolo, Jonathan. 2016. "News from the Other Side: How Topic Relevance Limits the Prevalence of Partisan Selective Exposure." *Journal of Politics* 78(3):763–773.

Supplementary Materials: Preferential Abstention in Conjoint Experiments

Contents

In the first portion of the Supplemental Materials (Section SMI) we report the popularity of forced-choice outcomes and the contexts for which they have been used in published conjoint experiments. Second, we build on our theoretical explanation of how abstention impacts the estimates of the AMCEs (Section SMII). We then discuss our replication of Funck and McCabe (2022) in Section SMIII, and Mummolo (2016) in Section SMIV. Specifically, we provide background information about the original survey experiments, and we present the full estimates from the models discussed in the manuscript.

SMI Popularity of Forced-Choice Outcomes in Conjoints

We surveyed the "top-3" journals in Political Science (American Journal of Political Science, American Political Science Review, and Journal of Politics) to capture all published research (articles and "letters"/"notes") that use a conjoint design in an experiment as of June 2022. We conducted a Google Scholar search with the following parameters: (a) Publication name–[the name of the journal in quotation marks], and (b) Keywords – "experiment" AND "discrete choice" OR "choice based" OR "stated preference" OR "conjoint". After saving the resulting list of articles, we retrieved the PDF file of each article to assess whether the study utilized original conjoint experiments.[1] We determined that 41 of the articles identified by our search terms utilized original conjoint experiments.[2]

We then examined each article to determine the 1) types of objects that respondents were asked to evaluate, and 2) type of outcome measure that the authors utilized to evaluate the results of their conjoint experiments. These pieces of information were often clearly identified at the beginning of a paper's methods section and often included a picture of a sample conjoint task. We categorized the types of objects that respondents evaluated into the following groups:

- Candidates for elected office (type of office)
- Policies (type of policy)

[1]One of the authors and a research assistant independently coded each article for whether it contained an original conjoint experiment, the types of objects respondents were asked to evaluate, and the type of outcome measure the authors included. In cases where the author and research assistant disagreed on any of these three points, the other author evaluated and resolved the conflict.

[2]We excluded from our count all publications whose conjoint experiments utilized only one profile per task. We do so because these experiments cannot, by design, incorporate a forced-choice outcome.

- Immigration/asylum applicants
- Other

The two most common outcome measures are "forced-choice", where respondents are required to choose one of the profiles presented, or "rating", where respondents are asked to rate each profile presented on an ordinal scale (i.e., very likely to vote for, somewhat likely, somewhat unlikely, very unlikely). A small number of studies also utilized both forced-choice and rating outcomes. Thus, studies' outcome measures were categorized into the following groups:

- Forced-choice
- Rating
- Forced-choice and rating

The results of our data collection are shown in Table SM.1. Of the 41 articles, 88% (36/41) utilized a forced-choice outcome, either exclusively (56%, or 23/41) or in conjunction with a rating outcome (31%, or 13/41). The remaining studies used only a rating outcome (5%, or 2/41), or a choice-based outcome that allowed respondents to choose an option other than one of the two profiles in the task and/or abstain (7%, or 3/41). The usage of forced-choice designs appears to be driven by studies focusing on candidates for elected office (18 of the 36 studies utilizing forced-choice outcomes). Therefore, the theoretical and empirical issues we outline may be particularly problematic for a large portion of published work using conjoint experiments.

Table SM.1: Conjoint experiments published in AJPS, APSR, and JOP in alphabetical order.

	Citation	Choice Object	Outcome Measure(s)
1	Auerbach and Thachil 2018	Other (slum leaders in India)	Forced-choice
2	Auerbach and Thachil 2020	Other (slum residents seeking help)	Forced-choice
3	Arias and Blair 2022	Immigration/asylum applicants	Forced-choice & rating
4	Bakker, Schumacher and Rooduijn 2021	Candidates for US House of Representatives	Forced-choice & rating
5	Ballard-Rosa, Martin and Scheve 2017	Policies (income tax)	Forced-choice
6	Bansak, Bechtel and Margalit 2021	Policies (austerity packages)	Forced-choice & rating
7	Barnett, A. Jamal and Monroe 2021	Other (job opportunities)	Forced-choice & rating

	Citation	Choice Object	Outcome Measure(s)
8	Blumenau and Lauderdale 2022	Persuasive arguments	Choice-based (allows respondents to say both options are equally persuasive)
9	Campbell et al. 2019	Candidates for British Parliament	Forced-choice
10	Carnes and Lupu 2016	Candidates for local office in Argentina, UK, and US	Forced-choice
11	Costa 2021	Candidates for US House of Representatives	Forced-choice
12	Dill and Schubiger 2021	Policies (use of military force)	Forced-choice & rating
13	Doherty, Dowling and Miller 2019	Candidates for US state legislatures	Forced-choice
14	Eggers, Vivyan and Wagner 2018	Candidates for British Parliament	Choice-based (allows for abstention or voting for a party not displayed)

	Citation	Choice Object	Outcome Measure(s)
15	Eshima and Smith 2022	Candidates for Japanese House of Councilors	Forced-choice
16	Frederiksen 2022	Candidates for president/prime minister in five democracies	Rating
17	Ghosn et al. 2021	Other (places to return to in Syria)	Choice-based with abstention option
18	Hainmueller and Hopkins 2015	Immigration applicants	Forced-choice & rating
19	Hankinson 2018	Other (urban development projects)	Forced-choice
20	Hanretty, Lauderdale and Vivyan 2020	Candidates for British Parliament	Forced-choice
21	Henderson et al. 2022	Candidates in US House of Representatives primaries	Forced-choice & rating
22	Johns and Kölln 2020	Candidates for elected office (parties)	Forced-choice & rating

	Citation	Choice Object	Outcome Measure(s)
23	Kennedy, Waggoner and Ward 2022	Policies (algorithms for forecasting criminal recidivism)	Forced-choice & rating
24	Levy 2022	Candidates for governor in Colombia	Forced-choice & rating
25	Magni 2022	Other (welfare recipients)	Forced-choice
26	Magni and Reynolds 2021	Candidates for national legislature in UK, US, and NZ	Forced-choice
27	Mummolo 2016	Other (news stories)	Forced-choice
28	Mummolo and Nall 2017	Other (communities to live in)	Forced-choice
29	Nelson and Witko 2021	Other (job offers)	Forced-choice & rating
30	Ono and Zilis 2022	Other (judges)	Forced-choice
31	Peterson 2017	Candidates for US House of Representatives	Forced-choice

	Citation	Choice Object	Outcome Measure(s)
32	Peterson and Simonovits 2018	Candidates for US House of Representatives	Forced-choice
33	Poertner 2021	Candidates for Bolivian Plurinational Legislative Assembly	Forced-choice
34	Rodon and Sanjaume-Calvet 2020	Policies (redistribution)	Rating
35	Schneider 2020	Candidates for EU Parliament	Forced-choice & rating
36	Spater 2022	Candidates for local office in India & Other (neighbors)	Forced-choice
37	Teele, Kalla and Rosenbluth 2018	Primary candidates for local, state, and national offices in the US	Forced-choice
38	Tellez 2021	Policies (peace agreements)	Forced-choice

	Citation	Choice Object	Outcome Measure(s)
39	Tsai, Trinh and Liu 2022	Candidates for township party secretaries	Forced-choice
40	Ward 2019	Other (immigrants to live in respondents' neighborhood)	Forced-choice & rating
41	Weaver 2021	Candidates for mayor in Peru	Forced-choice

Notes: This table presents information about the choice objects and outcome measures utilized in all conjoint experiments published in the *American Political Science Review*, *American Journal of Political Science*, and *Journal of Politics* as of June 2022. For more information about each of the experiments included in this table, please refer to the original articles.

SMII Theoretic Explanation of Abstention Options and AMCEs

Following from our manuscript, we consider the implications of including an abstention option versus a forced-choice design for conjoint experiments. We focus on the real-world context in which voters might not participate in an election if the only candidates they could select from have undesirable characteristics. We elaborate on our theoretical illustration more extensively in this section, specifically the precise implications to the AMCEs for both 5- and N-voter elections.

SMII.2 5-Voter Example

In our first scenario of selective abstention, we assume that voters uniformly turn out to vote and make a selection only if at least one of the candidates in the choice set possesses the characteristic they prioritize—pro-life (PL) for voters V1, V3, and V5, and pro-choice (PC) for voters V2 and V4. The five voters' aggregate preferences under uniform abstention are displayed in Table SM.2.

Note that in the final two comparisons (Table SM.2a, Rows 5 and 6), although all voters prefer the first candidate to the second candidate, those candidates receive fewer votes because some voters would abstain if presented with those comparisons. When we then calculate the AMCEs for the two attributes in Table SM.2b, we obtain different quantities than we would with no abstentions. In fact, our AMCEs moved in the opposite direction. The AMCE for pro-life has increased from 0.10 to 0.15, while the AMCE for corruption has decreased by 0.025 from 0.05 to 0.025. Even though the underlying preferences of the voters remained constant, the estimated AMCEs obtained are different.

Table SM.2: Scenario with uniform abstention.

(a) Individual and aggregate preferences

Comparison	V1	V2	V3	V4	V5	Tally
PL, D vs. PC, I	PL, D	PC, I	PL, D	PC, I	PL, D	3,2
PL, D vs. PC, D	PL, D	PC, D	PL, D	PC, D	PL, D	3,2
PL, D vs. PL, I	PL, D	-	PL, D	-	PL, D	3,0
PL, I vs. PC, I	PL, I	PC, I	PL, I	PC, I	PL, I	3,2
PL, I vs. PC, D	PL, I	PC, D	PL, I	PC, D	PL, I	3,2
PC, D vs. PC, I	-	PC, I	-	PC, I	-	0,2

(b) AMCEs for pro-life and cutting taxes

Panel A: Pro-life	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PC,D;PL,D)$	-0.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PC,D;PL,I)$	1
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PC,D;PC,D)$	2
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PC,D;PC,I)$	3
$\bar{Y}(PL,I;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	-2
$\bar{Y}(PL,I;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	-0.50
$\bar{Y}(PL,I;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	1
$\bar{Y}(PL,I:PC,I)$ - $\bar{Y}(PC,I;PC,I)$	2
	AMCE = $\frac{6}{40} = 0.15$
Panel B: Cutting taxes	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PL,I;PL,D)$	1.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PL,I;PL,I)$	1.50
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PL,I;PC,D)$	0
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PL,I;PC,I)$	0
$\bar{Y}(PC,D;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0
$\bar{Y}(PC,D;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0
$\bar{Y}(PC,D;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	-1
$\bar{Y}(PC,D;PC,I)$ - $\bar{Y}(PC,I;PC,I)$	-1
	AMCE = $\frac{1}{40} = 0.025$

Notes: PL=Pro-life, PC=Pro-choice, D=Decrease upper-class taxes, I=Increase upper-class taxes.

Second, consider what would happen if *only* one of the two types of voters chose not to participate when neither candidate had the level of the attribute the voter weighed more heavily in their decision-making. First, we can imagine the preference orderings if those emphasizing pro-life candidates abstained (Table SM.3a). This time the third and the sixth comparison are affected by voters' abstentions relative to when all voters choose a candidate. Again, when we calculate the AMCEs for the two attributes in Table SM.3b we see that our AMCEs are different; by abstaining, the pro-life voters have made the pro-life AMCE seem stronger and the effect of cutting taxes seem weaker. In fact, the AMCE for cutting taxes is now negative rather than positive, suggesting that candidates who pledge to raise taxes are now more preferred than those pledging to cut taxes *despite* the number of voters preferring tax-cutting candidates exceeding the number who prefer tax-raising candidates.

Next, we imagine what would happen if those emphasizing pro-choice candidates abstained in Table SM.4a. In this instance, only the last comparison is affected by voters' abstentions relative to when all voters choose a candidate. Now, when we calculate the AMCEs for the two attributes in Table SM.4b, we see that when voters prefer pro-choice candidates and abstain when only pro-life candidates are on the ballot, the AMCE for pro-life decreases to 0.00, but the AMCE for cutting taxes increases to 0.10.

The first takeaway is that allowing for abstentions can increase or decrease the AMCE estimates. Second, if abstention rates are even across voter types and the number of voters of each type is not equal, the AMCE for the attribute that is more important to the larger group becomes greater and the attribute that is more important to the smaller group gets smaller. Third, when only one type of voter abstains, the AMCE for the

Table SM.3: Scenario with pro-life abstention.

(a) Aggregate preferences over candidate profiles.

Comparison	V1	V2	V3	V4	V5	Tally
PL, D vs. PC, I	PL, D	PC, I	PL, D	PC, I	PL, D	3,2
PL, D vs. PC, D	PL, D	PC, D	PL, D	PC, D	PL, D	3,2
PL, D vs. PL, I	PL, D	PL, I	PL, D	PL, I	PL, D	3,2
PL, I vs. PC, I	PL, I	PC, I	PL, I	PC, I	PL, I	3,2
PL, I vs. PC, D	PL, I	PC, D	PL, I	PC, D	PL, I	3,2
PC, D vs. PC, I	-	PC, I	-	PC, I	-	0,2

(b) AMCEs for pro-life and cutting taxes

Panel A: Pro-life	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PC,D;PL,D)$	0.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PC,D;PL,I)$	1
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PC,D;PC,D)$	2
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PC,D;PC,I)$	3
$\bar{Y}(PL,I;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0
$\bar{Y}(PL,I;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0.50
$\bar{Y}(PL,I;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	1
$\bar{Y}(PL,I:PC,I)$ - $\bar{Y}(PC,I;PC,I)$	2
	AMCE = $\frac{10}{40} = 0.25$

Panel B: Cutting taxes	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PL,I;PL,D)$	0.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PL,I;PL,I)$	0.50
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PL,I;PC,D)$	0
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PL,I;PC,I)$	0
$\bar{Y}(PC,D;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0
$\bar{Y}(PC,D;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0
$\bar{Y}(PC,D;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	-1
$\bar{Y}(PC,D;PC,I)$ - $\bar{Y}(PC,I;PC,I)$	-1
	AMCE = $-\frac{1}{40} = -0.025$

Notes: PL=Pro-life, PC=Pro-choice, D=Decrease upper-class taxes, I=Increase upper-class taxes.

Table SM.4: Scenario with pro-choice abstention.

(a) Aggregate preferences over candidate profiles.

Comparison	V1	V2	V3	V4	V5	Tally
PL, D vs. PC, I	PL, D	PC, I	PL, D	PC, I	PL, D	3,2
PL, D vs. PC, D	PL, D	PC, D	PL, D	PC, D	PL, D	3,2
PL, D vs. PL, I	PL, D	-	PL, D	-	PL, D	3,0
PL, I vs. PC, I	PL, I	PC, I	PL, I	PC, I	PL, I	3,2
PL, I vs. PC, D	PL, I	PC, D	PL, I	PC, D	PL, I	3,2
PC, D vs. PC, I	PC, D	PC, I	PC, D	PC, I	PC, D	3,2

(b) AMCEs for pro-life and cutting taxes

Panel A: Pro-life	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PC,D;PL,D)$	-0.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PC,D;PL,I)$	1
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PC,D;PC,D)$	0.50
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PC,D;PC,I)$	0
$\bar{Y}(PL,I;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	-2
$\bar{Y}(PL,I;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	-0.50
$\bar{Y}(PL,I;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	1
$\bar{Y}(PL,I:PC,I)$ - $\bar{Y}(PC,I;PC,I)$	0.50
	AMCE = $\frac{0}{40} = 0.00$
Panel B: Cutting taxes	
$\bar{Y}(PL,D;PL,D)$ - $\bar{Y}(PL,I;PL,D)$	1.50
$\bar{Y}(PL,D;PL,I)$ - $\bar{Y}(PL,I;PL,I)$	1.50
$\bar{Y}(PL,D;PC,D)$ - $\bar{Y}(PL,I;PC,D)$	0
$\bar{Y}(PL,D;PC,I)$ - $\bar{Y}(PL,I;PC,I)$	0
$\bar{Y}(PC,D;PL,D)$ - $\bar{Y}(PC,I;PL,D)$	0
$\bar{Y}(PC,D;PL,I)$ - $\bar{Y}(PC,I;PL,I)$	0
$\bar{Y}(PC,D;PC,D)$ - $\bar{Y}(PC,I;PC,D)$	0.50
$\bar{Y}(PC,D;PC,I)$ - $\bar{Y}(PC,I;PC,I)$	0.50
	AMCE = $\frac{4}{40} = 0.10$

Notes: PL=Pro-life, PC=Pro-choice, D=Decrease upper-class taxes, I=Increase upper-class taxes.

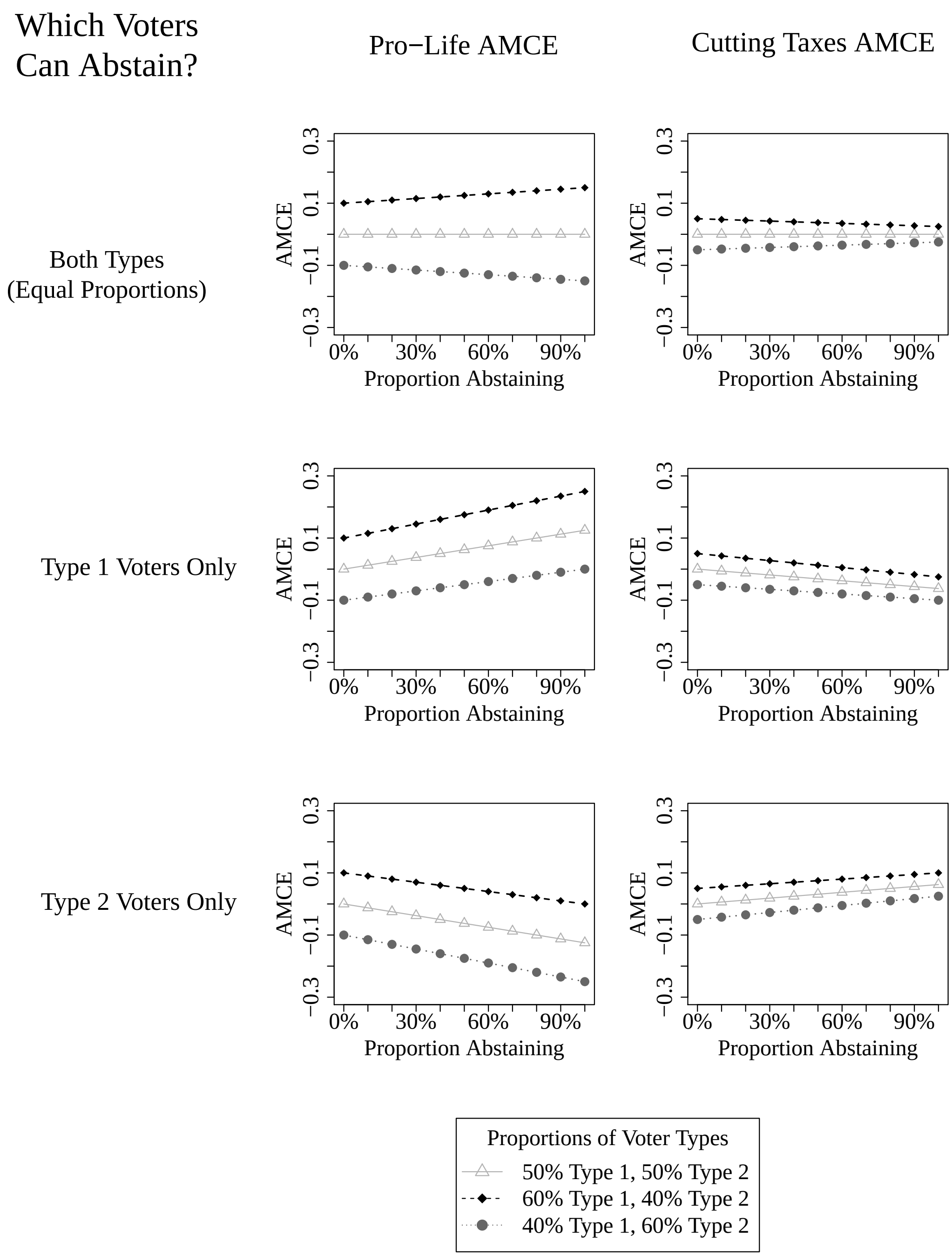
attribute that is more important for that group becomes greater, while the attribute that is more important to the other group becomes smaller (i.e. the abstaining group is no longer contributing "1"s to choices it would never choose, thus increasing the difference in marginal means for that attribute). To show how the rate of abstention among voters with different preference orderings impacts the AMCEs, we extend our example to N-voters.

SMII.3 Extension to N-Voters

In this extension, we imagine that voters are one of two types; the preference orderings of Type 1 voters mirror those of voters 1, 3, and and 5 in our 5-voter example and the preference orderings of Type 2 voters match those of voters 2 and 4. The rows of Figure SM.1 indicate which types of voters are allowed to abstain if no candidate with their prioritized level of the abortion policy attribute is present, and the columns differentiate the AMCEs for pro-life (first column) and cutting taxes (second column) under different rates of abstention identified by the x-axis of each pane. We also vary the proportions of Type 1 and Type 2 voters in the electorate, such that there can be an even number of Type 1 and Type 2 voters (light grey triangles), Type 1 voters outnumber Type 2 voters by a 3-to-1 margin (black diamonds), or Type 2 voters outnumber Type 1 voters by the same 3-to-1 margin (dark grey circles).

Given these conditions, we examine in Figure SM.1 how changes in the rate of abstention for one or both types of voters affects our AMCEs of interest. We observe in the first row of Figure SM.1 that when both types of voters abstain in equal proportions and the proportion of voters types is not equal, that the magnitude of the AMCE associated with the attribute prompting abstention increases as abstension increases.

Figure SM.1: Heterogeneous preferences for both attributes and identical prioritization.



Notes: The AMCEs that would be obtained using a forced-choice design are those presented in each pane where the proportion of voters abstaining is 0%.

Yet, the magnitude of the AMCE associated with an attribute that does not prompt abstention attenuates toward zero. Differently, in the second and third rows of Figure SM.1, only pro-life and pro-choice voters can abstain if neither candidate has their preferred position on abortion policy. It appears that as abstention increases, the AMCE for pro-life candidates moves toward the preferences of those abstaining, while the AMCE for candidates pledging to cut taxes moves away from the preferences of abstaining voters (e.g., in the second row, as more pro-life voters abstain, the AMCE for pro-life increases while the AMCE for cutting taxes decreases). Particularly, these shifts in AMCEs lead to differences in not only magnitude but also directionality, as the AMCE for cutting taxes changes signs when the voter type abstaining is more common and the abstention rate exceeds 70%. For instance, in the bottom right pane, when the voter composition is 40% Type 1 and 60% Type 2 and Type 2 voters abstain at a rate of greater than 70%, the AMCE for candidates pledging to cut taxes changes from negative to positive.

It is important to note that while this N-voter example accounts for shifts in the rate of abstention among different voter types, we have fixed two other facets of the conjoint setting. First, we have assumed that both types of voters have heterogeneous preferences over the two attributes; for each attribute, Type 1 and Type 2 voters prefer different levels (e.g., Type 1 voters prefer pro-life candidates, Type 2 voters prefer pro-choice candidates). Second, we have assumed that both types of voters have the same priorities over the included attributes; both Type 1 and Type 2 voters prioritize candidates' abortion policy positions over their tax policy positions. In additional simulations, we explore the implications that varying these facets can have on the resulting AMCEs. In brief, we find that the rate of abstention affects the magnitude of the AMCEs recovered and that sign-switching is possible across the range of abstention rates so long as both voter types do not have homogeneous preferences for all attributes.

Our theoretical exercise demonstrates that the estimated AMCEs can differ in both magntitude and sign depending on whether the choice-based outcome measure employed forces respondents to select a profile in each task. This is problematic because many contexts to which conjoint experiments are applied, such as voting for candidates for elected office, individuals may not ordinarily be forced to express a preference. As such, AMCEs based on forced-choice outcomes may often not reflect a sample's, nor a population's, aggregate preferences as expressed through real-world behavior.

SMIII Replication and Extension of Funck and McCabe (2022)

In their original article, Funck and McCabe (F-M) provide 2,135 American respondents recruited through Lucid with two hypothetical candidates competing for U.S. Congress in an upcoming general election.[3] In each task, one candidate was randomly assigned to be a Democrat and the other was assigned to be a Republican. After viewing the candidate profiles, respondents provide their vote choice for one candidate only. Participants completed three comparison tasks for a total of six candidate profiles per respondent, and an overall number of 12,810 candidate profiles.

To replicate F-M, we recruited a sample of 2,254 American respondents using

[3] Each recent news attribute-level (as well as those for candidates' policy stances, age, and religion) had an equal probability of appearing in every candidate profile and randomized across tasks as well as respondents, but the probability of appearance for some attribute-levels such as those concerning race, profession, and gender were adjusted to reflect the probability distribution of candidates in the real-world.

Amazon's Mechanical Turk (MTurk). The design of our replication (question wording, probability distribution of attribute-levels, etc.) mirrored exactly that of F-M with the following three exceptions:

1. As we describe in the main text, respondents were randomly assigned to one of two experimental arms—a *forced-choice arm*, where they were required to indicate which candidate they preferred in each task, or a *abstention option arm*, where they could indicate their preferred candidate or abstain from making a selection.

2. Respondents completed six comparison tasks in their assigned experimental arm for a total of twelve candidate profiles per respondent, and an overall number of 26,320.[4]

3. After completing their six comparison tasks, we asked respondents to complete a final, seventh task in which they were provided with the outcome measure from the opposite treatment arm. Our analyses in the main paper utilize only responses provided in the first six comparison tasks completed by each respondent, and we examine respondents' choices in this seventh task separately in this section.

To assess how the AMCEs differ across the forced-choice and abstention arms, we use ordinary least squares regression to estimate our binary indicator of whether a respondent selected a profile in a given task as a function of a triple interaction consisting of the following terms:

[4]We asked respondents to complete six comparison tasks, unlike the three F-M asked respondents to complete, for power considerations; because we wanted to compare AMCEs across experimental arms, doubling the number of tasks completed by a similar number of respondents yields a sample in each treatment arm similar in size to the full sample used by F-M.

- A binary indicator for whether the respondent was in the forced-choice (0) or abstention (1) arm.
- A series of binary indicators for each attribute-level of information environment complexity (excluding the reference category, Low).
- A series of binary indicators for each attribute-level of recent news (excluding the reference category, No recent news).

Table SM.5 reports the estimated coefficients this model, and Figure SM.2 presents visual representations of the AMCEs and the pairwise differences of each AMCE among respondents who were in the forced-choice and abstention option arms. One of the fifteen pairwise differences is distinguishable from zero—the difference in the AMCEs for Sexual Harassment in the Hign Information condition in the forced-choice and abstention arms. Substantively, this means that when respondents in the High Information condition were able to abstain from choosing a candidate, they were less likely to select candidates implicated in a sexual harassment scandal relative to respondents in the abstention arm.

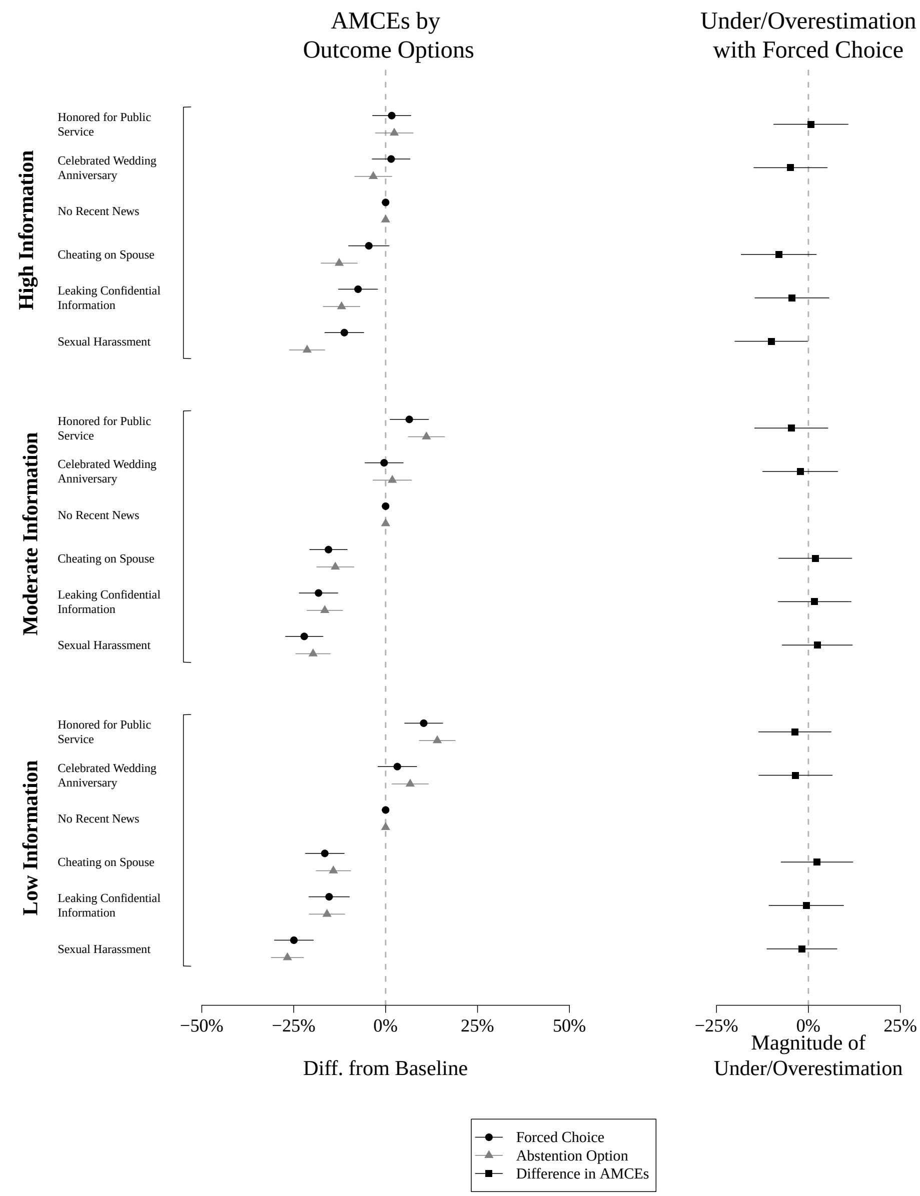
However, it is important to emphasize that the hypotheses of F-M focus on the *differences* in AMCEs across information environment facets rather than the AMCEs for each attribute-level in each information environment. Accordingly, Figure 1 in the manuscript, uses the regression model reported in Table SM.5 to calculate the difference between the AMCEs for each non-baseline attribute-level when situated in a low information environment conjoint task as compared to when situated in a moderate or high information environment conjoint task.

Table SM.5: Replication results testing "Information Hypothesis" of F-M.

	Forced		Abstention	
	Ref. Cat. = Low	Ref. Cat. = Medium	Ref. Cat. = Low	Ref. Cat. = Medium
Intercept	0.57*	0.58*	0.43*	0.44*
	(0.02)	(0.02)	(0.02)	(0.02)
Medium information	0.01		0.01	
	(0.02)		(0.02)	
High information	−0.04	−0.05*	0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Cheating on spouse	−0.17*	−0.16*	−0.14*	−0.14*
	(0.03)	(0.03)	(0.02)	(0.03)
Leaking confidential information	−0.15*	−0.18*	−0.16*	−0.17*
	(0.03)	(0.03)	(0.02)	(0.02)
Sexual harassment	−0.25*	−0.22*	−0.27*	−0.20*
	(0.03)	(0.03)	(0.02)	(0.02)
Wedding anniversary	0.03	−0.00	0.07*	0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Honored public service	0.10*	0.06*	0.14*	0.11*
	(0.03)	(0.03)	(0.02)	(0.03)
Medium information x cheating on spouse	0.01		0.01	
	(0.04)		(0.03)	
Medium information x leaking confidential information	−0.03		−0.01	
	(0.04)		(0.03)	
Medium information x sexual harassment	0.03		0.07*	
	(0.04)		(0.03)	
Medium information x wedding anniversary	−0.04		−0.05	
	(0.04)		(0.04)	
Medium information x honored public service	−0.04		−0.03	
	(0.04)		(0.03)	
High information x cheating on spouse	0.12*	0.11*	0.02	0.01
	(0.04)	(0.04)	(0.03)	(0.04)
High information x leaking confidential information	0.08*	0.11*	0.04	0.05
	(0.04)	(0.04)	(0.03)	(0.04)
High information x sexual harassment	0.14*	0.11*	0.05	−0.02
	(0.04)	(0.04)	(0.03)	(0.03)
High information x wedding anniversary	−0.02	0.02	−0.10*	−0.05
	(0.04)	(0.04)	(0.04)	(0.04)
High information x honored public service	−0.09*	−0.05	−0.12*	−0.09*
	(0.04)	(0.04)	(0.04)	(0.04)
N	12620	8366	13700	9080

Notes: We mirror Table 1 from F-M by estimating the same regression models using our experimental data but estimating separate models for the forced-choice and abstention arms. Standard errors are clustered by respondent, shown in parentheses, and statistical significance is indicated as $^{*}p < 0.05$.

Figure SM.2: AMCEs for the forced-choice and abstention-option experimental arms replicating the design of F-M.



Notes: The left pane presents, by quantity of information, the average marginal component effect of each characteristic (type of scandal) on the predicted likelihood that respondents select a given candidate. The AMCEs and their corresponding 95% confidence intervals are represented by the horizontal lines. The right pane presents the differences in the magnitudes of the AMCEs for each attribute-level between the forced-choice and abstention option designs (i.e. $|AMCE_{\text{Forced}} - AMCE_{\text{Abstention}}|$) and Bonferroni-adjusted 95% confidence intervals ($\alpha = \frac{0.05}{15} = 0.00\overline{3}$).

SMIII.2 Alternative Coding of Abstentions

In the replication analyses we present in the main paper, we code the outcome measure for each profile in which a respondent abstained, or chose none of the available profiles, as “0.” This coding practice follows the formal presentation of the AMCE by Hainmueller, Hopkins and Yamamoto, where the outcome $Y_{ijk}(\bar{\mathbf{t}})$ “is a binary indicator variable where 1 indicates that respondent i would choose the jth profile in her kth task if she got the treatment set $\bar{\mathbf{t}}$ and 0 implies that she would not” (2014, 7), and the AMCE is calculated as “the difference in the average observed choice outcomes between the treatment ($T_{ijkl} = t_1$) and control ($T_{ijkl} = t_0$) groups within each stratum” (2014, 14).

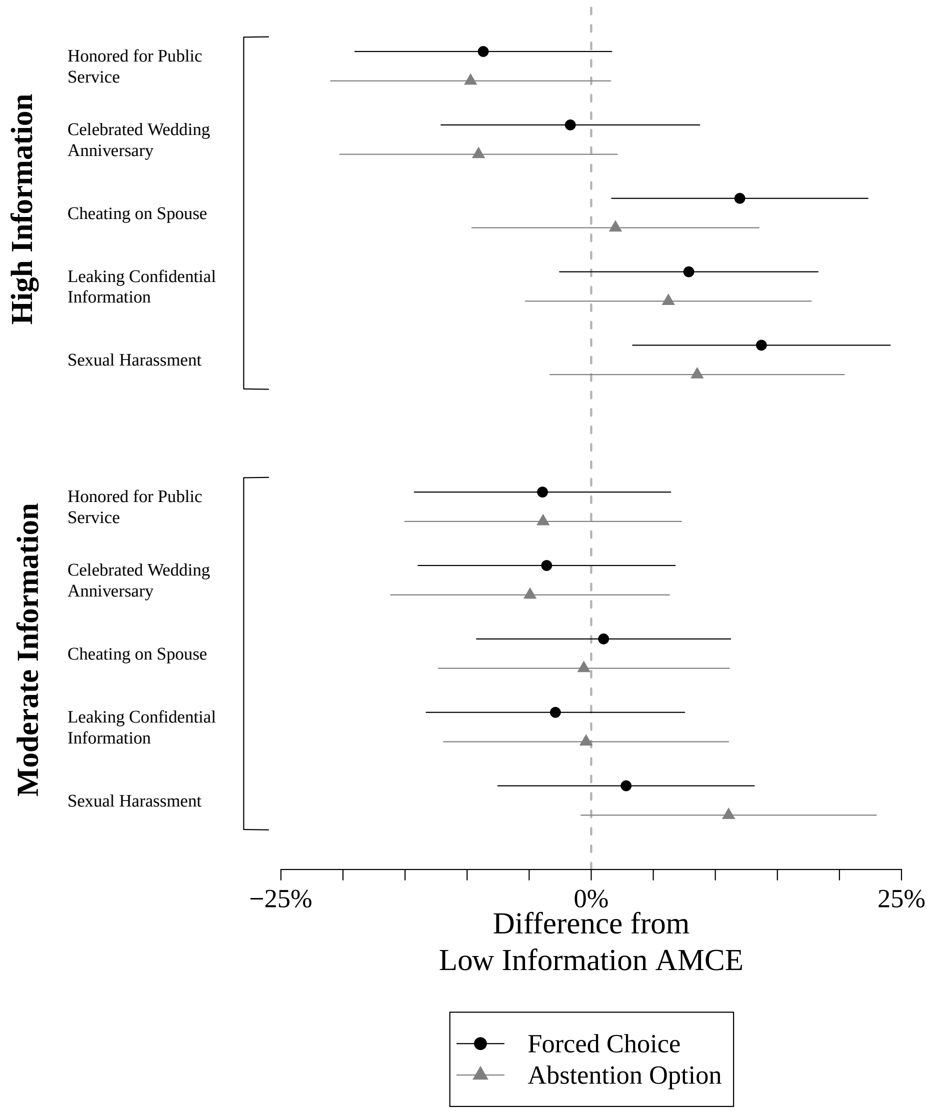
However, if a researcher wanted to treat abstentions as substantively different responses rather than merely the absence of choosing one of the available profiles, respondents who abstain may alternatively be coded as missing (i.e., NA). To assess how this might affect AMCEs, we repeated our analyses by coding the choice outcome for each profile in which a respondent abstained as missing (i.e., dropping it from the analysis). The results of this alternative specification is presented in Table SM.6 and Figure SM.3. Compared to their analogues which code abstentions as zeroes (Table SM.5 and Figure 1, respectively), the estimates are substantively similar, though the confidence intervals for the differences between the "Low Information" AMCEs and the "Medium" and "High Information" AMCEs are wider when abstentions are coded as missing. Thus, in the context of this replication, coding abstention outcomes as missing rather than zero does not affect the substantive interpretation of the results.

Table SM.6: Replication results testing "Information Hypothesis" of F-M, alternative outcome coding.

	Forced		Abstention	
	Ref. Cat. = Low	Ref. Cat. = Medium	Ref. Cat. = Low	Ref. Cat. = Medium
Intercept	0.57*	0.58*	0.57*	0.57*
	(0.02)	(0.02)	(0.02)	(0.02)
Medium information	0.01		0.00	
	(0.02)		(0.03)	
High information	−0.04	−0.05*	0.01	0.01
	(0.02)	(0.02)	(0.03)	(0.03)
Cheating on spouse	−0.17*	−0.16*	−0.15*	−0.15*
	(0.03)	(0.03)	(0.03)	(0.03)
Leaking confidential information	−0.15*	−0.18*	−0.18*	−0.19*
	(0.03)	(0.03)	(0.03)	(0.03)
Sexual harassment	−0.25*	−0.22*	−0.31*	−0.20*
	(0.03)	(0.03)	(0.03)	(0.03)
Wedding anniversary	0.03	−0.00	0.05	−0.00
	(0.03)	(0.03)	(0.03)	(0.03)
Honored public service	0.10*	0.06*	0.12*	0.08*
	(0.03)	(0.03)	(0.03)	(0.03)
Medium information x cheating on spouse	0.01		−0.01	
	(0.04)		(0.04)	
Medium information x leaking confidential information	−0.03		−0.00	
	(0.04)		(0.04)	
Medium information x sexual harassment	0.03		0.11*	
	(0.04)		(0.04)	
Medium information x wedding anniversary	−0.04		−0.05	
	(0.04)		(0.04)	
Medium information x honored public service	−0.04		−0.04	
	(0.04)		(0.04)	
High information x cheating on spouse	0.12*	0.11*	0.02	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
High information x leaking confidential information	0.08*	0.11*	0.06	0.07
	(0.04)	(0.04)	(0.04)	(0.04)
High information x sexual harassment	0.14*	0.11*	0.09*	−0.03
	(0.04)	(0.04)	(0.04)	(0.04)
High information x wedding anniversary	−0.02	0.02	−0.09*	−0.04
	(0.04)	(0.04)	(0.04)	(0.04)
High information x honored public service	−0.09*	−0.05	−0.10*	−0.06
	(0.04)	(0.04)	(0.04)	(0.04)
N	12620	8366	10172	6778

Notes: We mirror Table 1 from F-M by estimating the same regression models using our experimental data but estimating separate models for the forced-choice and abstention arms. Unlike in our main analysis, outcomes for profiles where respondents in the abstention arm chose to abstain are coded as missing. Standard errors are clustered by respondent, shown in parentheses, and statistical significance is indicated as $^{*}p < 0.05$.

Figure SM.3: Effect of scandalous news and information complexity on vote choice, alternative outcome coding.



Notes: This figure presents the differences in the AMCE for each non-baseline "Recent News" attribute-level for conjoint tasks situated in a "Low Information" conjoint task relative to when the same attribute-level is situated in "Moderate" (lower half of figure) or "High Information" (upper half of figure). The black circles and lines represent differences and Bonferroni-adjusted 95% confidence intervals ($\alpha = \frac{0.05}{20} = 0.0025$) for each attribute-level among respondents in the forced-choice arm (Liu and Shiraito 2023). The grey triangles and lines represent those same quantities among respondents in the abstention option arm. Unlike in our main analysis, outcomes for profiles where respondents in the abstention arm chose to abstain are coded as missing. See Table SM.6 for the full model summary.

SMIII.3 When Do Respondents Abstain?

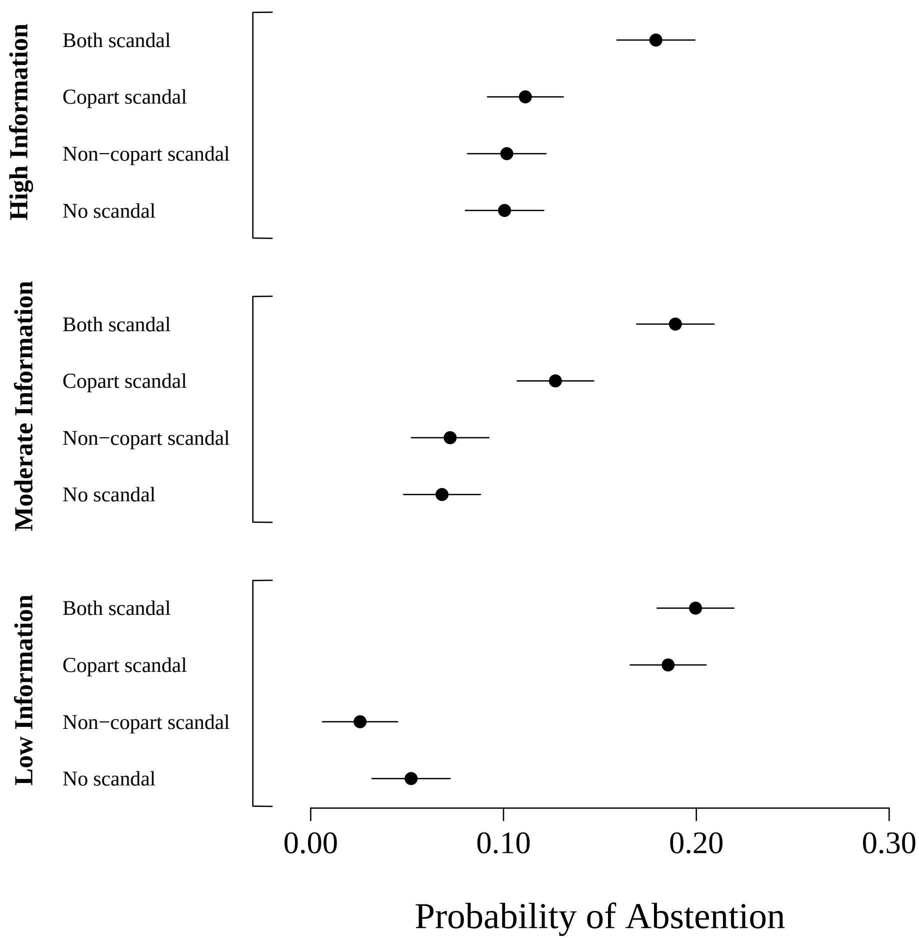
Among respondents in our abstention arm, 31.0% (699 of 2254) chose to abstain in at least one task, and the abstention option was selected in 25.8% (1764 of 6850) of all tasks completed by those in the abstention arm. To better understand why respondents in the abstention arm chose to abstain in some tasks, we used ordinary least squares to regress a binary indicator for whether the respondent abstained in a given task (1) or not (0) on an interaction between the following two task and profile characteristics:

- A series of binary indicators for which candidates had Recent News attribute-levels implicating them in scandals—Neither candidate (reference category), only the candidate who shared the respondent's partisan affiliation, only the candidate who did not share the respondent's partisan affiliation, or both candidates.
- A series of binary indicators for whether the task was situated in a low (reference category), medium, or high information environment.

Figure SM.4 indicates the predicted probability that a respondent would abstain from a task given its information environment complexity as well as whether and which candidates are implicated in scandals. The relative magnitudes of these predicted probabilities provide several insights concerning when respondents are more likely to abstain.

First, we can see that across information environments, respondents are most likely to abstain when both candidates experience scandal. This is consistent with the notion that respondents would rather abstain than endorse a candidate with negative valence. Second, respondents are more likely to abstain when only the candidate who shares

Figure SM.4: Predicted probabilities of abstention in F-M replication given task characteristics.



Notes: The figure plots the predicted probability that a respondent will abstain from choosing either candidate in a task given the complexity of the task's information environment and whether and which candidates have experienced scandal.

their partisan affiliation experiences scandal in comparison to when only the candidate who does not share their partisan affiliation or neither candidate experiences scandal. This is consistent with the notion that respondents would rather abstain than choose a candidate who does not share their partisan affiliation. However, this difference in the probabilities of abstention shrink when moving from the low to moderate and high information environments, which suggests that respondents might use additional information about copartisan candidates to rationalize away their scandals.

Third, respondents are more likely to abstain as the complexity of the information environment increases. This may indicate that some respondents would rather abstain than exert the cognitive effort to consider higher levels of detail. Conversely, as more information about candidates is presented, it becomes more likely that each candidate would have at least one undesirable level of an attribute, which prompts respondents to abstain.

SMIII.4 Who Abstains?

Following extant literature on the individual-level characteristics associated with abstention and nonresponse in contexts such as voting and responding to survey questions, like gender, race, and levels of education and income (e.g., Berinsky 2008; Rosenstone and Hansen 2003), we also assess whether those same characteristics are related to abstention in our replication. Before completing any of our conjoint tasks, we asked respondents to provide information about the following demographic characteristics:

- Age (four-point ordinal scale)
- Gender (1 if female, 0 otherwise)
- Educational attainment (five-point ordinal scale)

- Race/ethnicity (binary indicators for Asian, Black, other non-white race, and Hispanic)
- Income (twelve-point ordinal scale)
- Ideology (seven-point ordinal scale)
- Party identification (seven-point ordinal scale)

Using this demographic information, we again used ordinary least squares to regress a binary indicator for whether the respondent abstained in a given task (1) or not (0) on each of these characteristics except party identification.[5] The estimates from this analysis are presented in Table SM.7. Some of the demographic characteristics generally thought to be related to abstention behave as expected; for instance, respondents identifying as female are 4 percentage points more likely to abstain than those identifying as male or non-binary and respondents who identify as Asian or with another race other than Black or White are 8 and 17 percentage points more likely to abstain, respectively. However, a few demographic characteristics are related to abstention in our replication in ways that differ from how they are commonly associated with abstention in other settings; for instance, Hispanic respondents are 11 percentage points less likely to abstain and moving up one point on the four-point age scale is associated with a 3 percentage point increase in abstaining.

Overall, these results highlight two important findings concerning abstention in conjoint experiments. First, as in other contexts of interest to political scientists,

[5]We omit party identification from this analysis because it is implicitly tied to the coding of task-level characteristics related to abstention in Section SMIII.2.

Table SM.7: Demographic correlates of abstention in F-M replication

	Model 1
Intercept	0.14^{*}
	(0.03)
Age	0.03^{*}
	(0.01)
Female	0.04^{*}
	(0.01)
Education	-0.01^{*}
	(0.01)
Asian	0.08^{*}
	(0.02)
Black	-0.00
	(0.02)
Other Non-White Race	0.17^{*}
	(0.03)
Hispanic	-0.11^{*}
	(0.02)
Income	0.00^{*}
	(0.00)
Ideology	0.02^{*}
	(0.00)
N	6645

Notes: Statistical significance is indicated as $^{*}p < 0.05$.

individual-level characteristics are indicative of abstention, and, to the extent that those characteristics have heterogeneous effects on some attribute-levels of interest, the estimates associated with a conjoint experiment may differ systematically depending on whether an abstention option is offered or withheld. Second, the associations between specific demographic characteristics and abstention in conjoint experiments may not perfectly mirror the relationships observed in other survey settings. So, researchers may have difficulty anticipating which respondents with which demographic characteristics are more inclined to abstain from choosing in a conjoint task if given the opportunity.

SMIII.5 Switching Experimental Arms

After completing the six tasks in their assigned experimental arm, respondents were asked to complete a seventh, final task in which the outcome measure reflected that used in the opposite experimental arm (e.g., respondents in the forced-choice arm were now given the opportunity to choose one of the candidates or to abstain). For each respondent, the attribute-levels of the two profiles and the information environment complexity for this seventh task were the same as shown in the first task, which allows us to assess how the same respondents evaluate the same set of profiles across the two different types of outcome measures.

Respondents in the abstention option arm were forced to choose one of the two candidates in this final task, even if they originally abstained; hence, the choices of abstention option arm respondents in this final task is deterministic (i.e., respondents who originally abstained must now choose one of the two profiles). However, respondents in the forced-choice arm provide a unique window on whether the same respondents would make a different choice if allowed to abstain relative to when they are forced to choose. Of the 1,035 respondents in the forced-choice arm who completed all seven tasks (the six forced-choice tasks and the seventh abstention option task), 286 (27.6%) chose to abstain rather than choosing the candidate they originally selected when forced to do so. This descriptive evidence suggests that a fairly large proportion of respondents will choose candidates if forced to do so even though they would ultimately prefer not to make a choice.

SMIV Replication and Extension of Mummulo (2016)

In Mummolo's original article, 1,059 American respondents that identified as Democrats or Republicans were recruited by Survey Sampling International (SSI). Participants were presented with pairs of news items, and in each task, the source and headline attributes of the news items are randomly assigned. Respondents were asked to indicate which news item they would prefer to read. Participants completed twelve comparison tasks for a total of twenty-four profiles per respondent, and an overall number of 25,416 profiles.

The dimensions of the news stories that Mummolo randomly varied concern the content and the source of the news stories, specifically their "relevance" (e.g. whether a participant was a member of an "affected population") and "friendliness", which is a partisan-based comparison between respondent (Democrat or Republican) and source (MSNBC, USA Today, or Fox News). Mummolo hypothesized and showed that respondents, on average, prefer content that is relevant and friendly, though "source reputations appear to do little to prevent consumption when topics are relevant" (2016:

771).

To replicate Mummolo, we recruited a sample of 1,553 American respondents who identified as Democrats or Republicans using Amazon's Mechanical Turk (MTurk). As in our replication of F-M, we randomly assigned respondents to a *forced-choice* or *abstention option* arm. Additionally, we again asked respondents to complete a final, thirteenth task in which they were provided with the outcome measure from the opposite treatment arm. Besides these two modifications, our replication otherwise mirrored the protocol used by Mummolo.

To assess how the AMCEs differ across the forced-choice and abstention arms, we use ordinary least squares regression with standard errors clustered by respondent. We estimate our binary indicator of whether a respondent selected a profile in a given task as a function of a triple interaction consisting of the following terms:

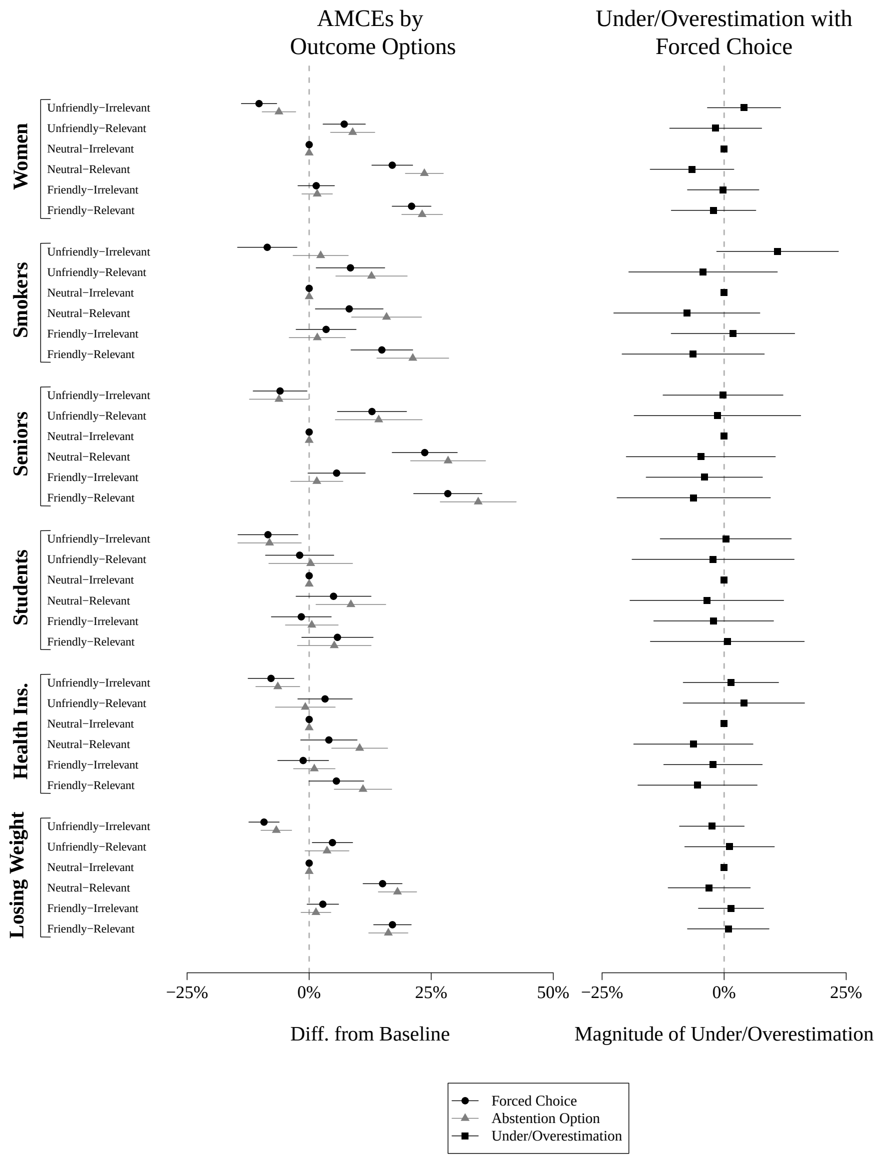
- A binary indicator for whether the respondent was in the forced-choice (0) or abstention (1) arm.
- A binary indicator for whether the news source was friendly, neutral, or unfriendly given the respondent's partisan affiliation.
- A binary indicator for whether the topic of the news item is relevant to the respondent, given the respondent's membership in one of six affected publics: women, smokers, senior citizens, college students, uninsured persons or healthcare workers, and those currently trying to lose weight.

Following Mummolo, we estimate separate models for respondents who are members of each affected public. AMCEs for respondents in the forced-choice and abstention option arms are shown with black circles and grey triangles, respectively, in the left pane of Figure SM.5, and the differences in the magnitudes of the AMCEs across arms. Additionally, in Table SM.8, we adapt Table 4 in Mummolo (2016, 770) by estimating separate models for each affected public and experimental arm.

The central finding from Mummolo is largely replicated: respondents in affected publics are more likely to select news items that are on topics relevant to them and do not come from unfriendly sources. While the AMCEs from the abstention option arm exhibit some differences in direction and magnitude compared to those in the forced-choice arm, none are statistically distinguishable at the Bonferroni-adjusted 95% confidence level.

However, some attribute-level-specific hypothesis tests yield different conclusions across treatment arms. For instance, the abstention option arm AMCEs for the Neutral-Relevant attribute-level among students and the uninsured and healthcare workers are statistically distinguishable from zero at the 95% level, but the same AMCEs from the forced-choice arm are not. In these cases, the topic relevance hypothesis presented by Mummolo 2016 enjoys more support when abstentions are allowed than when choices are forced, as in the original paper.

Figure SM.5: AMCEs for the forced-choice and abstention-option experimental arms replicating the design of Mummolo (2016).



Notes: The left pane presents, by affected population, the average marginal component effect of each non-baseline attribute-level (topic relevance and source friendliness) on the probability that respondents select a given story. The AMCEs and their corresponding 95% confidence intervals are represented by the horizontal lines. The right pane presents the differences in the magnitudes of the AMCEs for each attribute-level between the forced-choice and abstention option designs (i.e. $|AMCE_{\text{Forced}} - AMCE_{\text{Abstention}}|$) and Bonferroni-adjusted 95% confidence intervals ($\alpha = \frac{0.05}{30} = 0.001\overline{6}$).

Table SM.8: Adaptation of Table 4 of Mummolo 2016 by experimental arm.

	Women		Smokers		Seniors		Students		Health Ins.		Losing Weight	
	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention
Intercept	0.46*	0.32*	0.47*	0.35*	0.43*	0.28*	0.51*	0.42*	0.50*	0.38*	0.48*	0.38*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Relevant Topic	0.17*	0.24*	0.08*	0.16*	0.24*	0.28*	0.05	0.09*	0.04	0.10*	0.15*	0.18*
	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)
Friendly Source	0.01	0.02	0.03	0.02	0.06	0.02	−0.02	0.01	−0.01	0.01	0.03	0.01
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Unfriendly Source	−0.10*	−0.06*	−0.09*	0.02	−0.06*	−0.06*	−0.08*	−0.08*	−0.08*	−0.06*	−0.09*	−0.07*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Relevant x Friendly	0.03	−0.02	0.03	0.04	−0.01	0.05	0.02	−0.04	0.03	−0.00	−0.01	−0.03
	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Relevant x Unfriendly	0.00	−0.08*	0.09	−0.05	−0.05	−0.08	0.01	−0.00	0.07	−0.05	−0.01	−0.08*
	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
N	8424	8528	2903	2796	3123	2276	2662	2586	4320	4539	10078	9647

Notes: We mirror Table 4 from Mummolo (2016: 770) by estimating the same regression models using our experimental data but estimating separate models for the forced-choice and abstention option arms. All standard errors are clustered on respondent. Statistical significance is indicated as $^{*}p < 0.05$.

SMIV.2 Alternative Coding of Abstentions

As we mentioned in Section SMIII.2, if a researcher wanted to treat abstentions as substantively different responses rather than merely the absence of choosing one of the available profiles, she may alternatively choose to code choice outcomes where respondents abstain as missing (i.e., NA). To assess how this might affect AMCEs, we repeated our analyses by coding the choice outcome for each profile in which a respondent abstained as missing (i.e., dropping it from the analysis). The results of this alternative specification is presented in Table SM.9 and Figure SM.6. Compared to their analogues which code abstentions as zeroes (Table SM.9 and Figure SM.5, respectively), the estimates are substantively similar, though the confidence intervals for the pairwise differences between the forced choice and abstention AMCEs are slightly larger. Thus, in the context of this replication, coding abstention outcomes as missing rather than zero does not affect the substantive interpretation of the results.

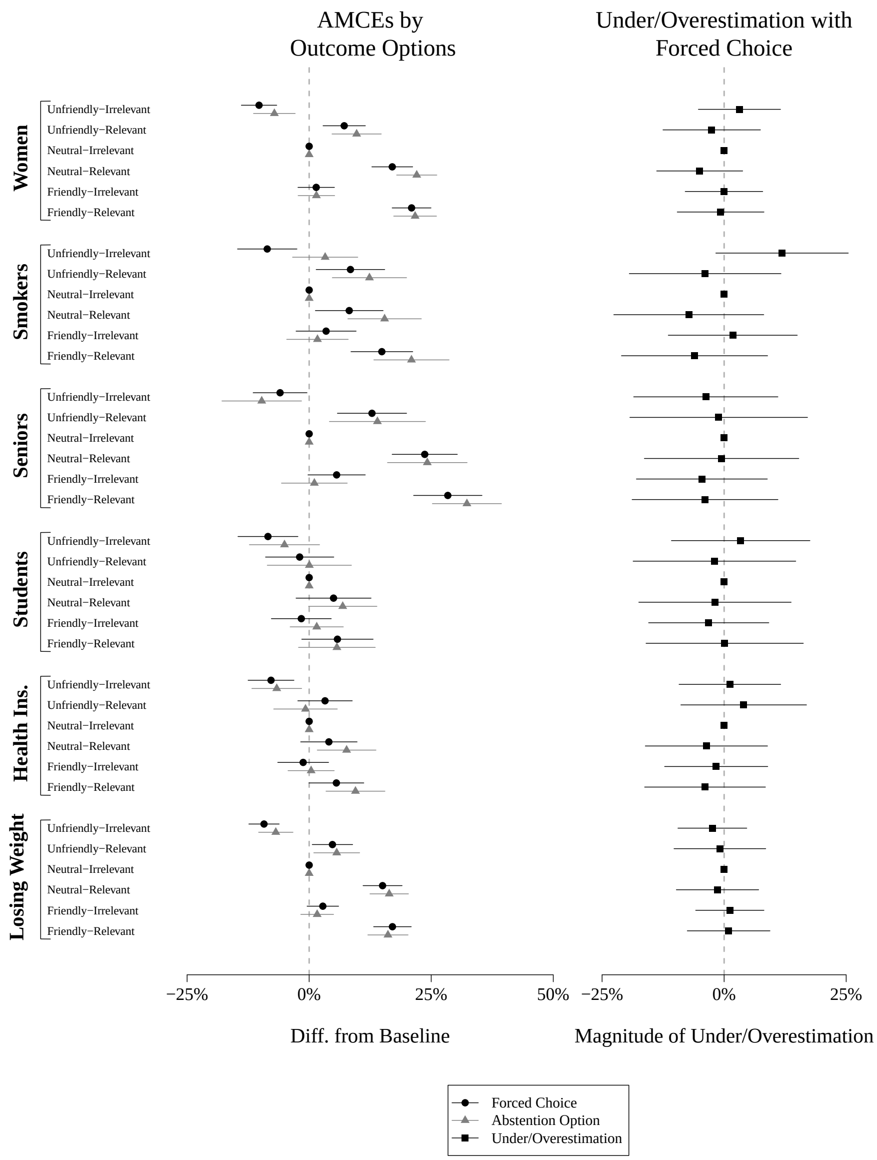
Rather than removing respondents from a sample, we encourage researchers to follow Häusermann, Kurer and Traber by controlling for socio-demographic factors that may contribute to abstention (2019, 1071) and comparing a subset of initial comparisons rather than all pairwise tasks due to respondent fatigue (2019, 1083).

Table SM.9: Adaptation of Table 4 of Mummolo 2016 by experimental arm, alternative outcome coding.

	Women		Smokers		Seniors		Students		Health Ins.		Losing Weight	
	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention	Forced	Abstention
Intercept	0.46*	0.43*	0.47*	0.43*	0.43*	0.42*	0.51*	0.49*	0.50*	0.49*	0.48*	0.47*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Relevant Topic	0.17*	0.22*	0.08*	0.15*	0.24*	0.24*	0.05	0.07	0.04	0.08*	0.15*	0.16*
	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)
Friendly Source	0.01	0.01	0.03	0.02	0.06	0.01	−0.02	0.02	−0.01	0.00	0.03	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Unfriendly Source	−0.10*	−0.07*	−0.09*	0.03	−0.06*	−0.10*	−0.08*	−0.05	−0.08*	−0.07*	−0.09*	−0.07*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	(0.03)	(0.02)	(0.02)
Relevant x Friendly	0.03	−0.02	0.03	0.04	−0.01	0.07	0.02	−0.03	0.03	0.01	−0.01	−0.02
	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Relevant x Unfriendly	0.00	−0.05	0.09	−0.06	−0.05	−0.00	0.01	−0.02	0.07	−0.02	−0.01	−0.04
	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)	(0.03)	(0.03)
N	8424	6532	2903	2366	3123	1629	2662	2178	4320	3551	10078	7901

Notes: We mirror Table 4 from Mummolo (2016: 770) by estimating the same regression models using our experimental data but estimating separate models for the forced-choice and abstention option arms. Unlike in our main analysis, outcomes for profiles where respondents in the abstention arm chose to abstain are coded as missing. All standard errors are clustered on respondent. Statistical significance is indicated as $^* p < 0.05$.

Figure SM.6: AMCEs for the forced-choice and abstention-option experimental arms replicating the design of Mummolo (2016), alternative outcome coding.



Notes: The left pane presents, by affected population, the average marginal component effect of each non-baseline attribute-level (topic relevance and source friendliness) on the probability that respondents select a given story. The AMCEs and their corresponding 95% confidence intervals are represented by the horizontal lines. The right pane presents the differences in the magnitudes of the AMCEs for each attribute-level between the forced-choice and abstention option designs (i.e. $|AMCE_{\text{Forced}} - AMCE_{\text{Abstention}}|$) and Bonferroni-adjusted 95% confidence intervals ($\alpha = \frac{0.05}{30} = 0.001\overline{6}$). Unlike in our main analysis, outcomes for profiles where respondents in the abstention arm chose to abstain are coded as missing.

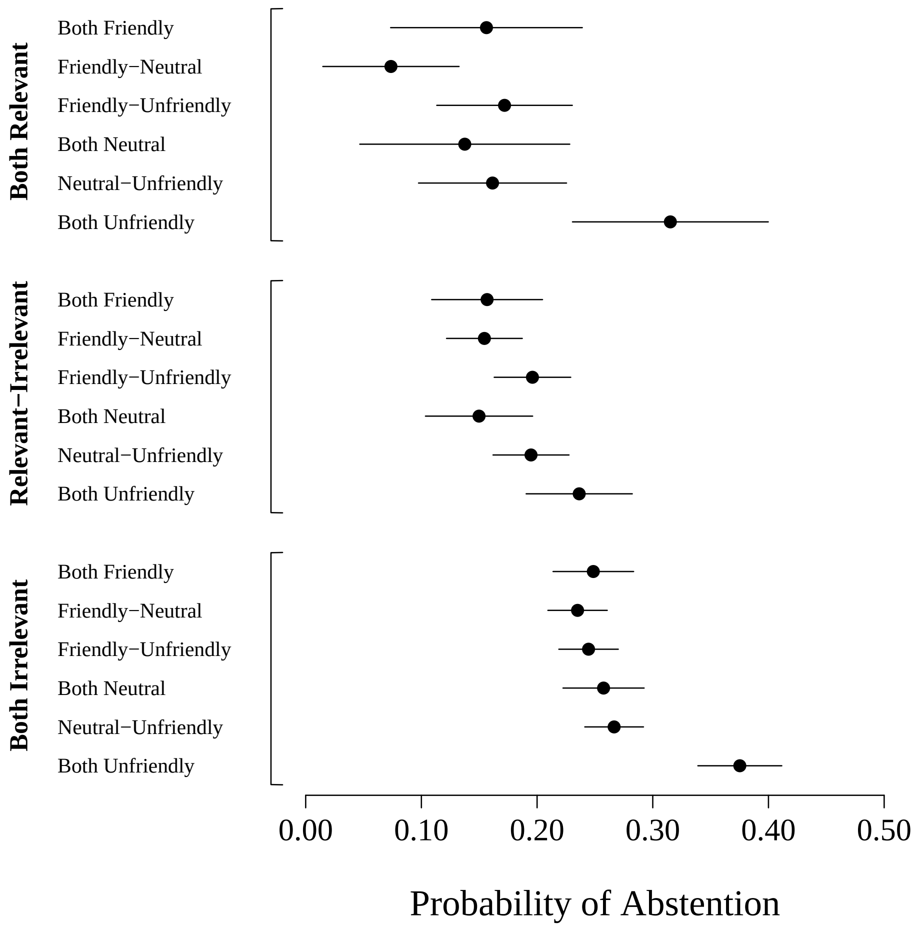
SMIV.3 When Do Respondents Abstain?

Among respondents in our abstention arm, 31.1% (565 of 1815) chose to abstain in at least one task, and the abstention option was selected in 24.3% (2335 of 9622) of all tasks completed by those in the abstention arm. To better understand why respondents in our abstention arm in some tasks, we used ordinary least squares to regress a binary indicator for whether the respondent abstained in a given task (1) or not (0) on an interaction between the following two task and profile characteristics:

- A series of binary indicators for the combination of the news items' topic relevance: whether both news items were irrelevant to the respondent (reference category); whether one news item was irrelevant and one was relevant; or whether both news items were relevant.
- A series of binary indicators for the combination of the news items' source friendliness: whether both news items came from an unfriendly source (reference category); whether one news item came from a neutral source and one news item came from an unfriendly source; whether both news items came from a neutral source; whether one news item came from a friendly source and one news item came from an unfriendly source; whether one news item came from a friendly source and one news item came from a neutral source; or whether both news items came from a friendly source.

Figure SM.7 indicates the predicted probability that a respondent would abstain from a task given the topic relevance and source friendliness of the presented stories. The relative magnitudes of these predicted probabilities provide several insights concerning

Figure SM.7: Predicted probabilities of abstention in Mummolo replication.



Notes: The figure plots the predicted probability that a respondent will abstain from choosing either news item in a task given the topic relevance and source friendliness of the presented news items.

when respondents are more likely to abstain. First, when focusing on topic relevance (the three bracketed sets), we see that respondents are most likely to abstain when both news items are irrelevant relative to when at least one of the news items is relevant. This is consistent with the notion that if presented only with news items that do not have respondents' optimal level of topic relevance, they would rather abstain than select a news item on an irrelevant topic. Second, when focusing on source friendliness (the six point estimates in each bracketed set), we see that respondents are most likely to abstain when the sources for both news items are unfriendly relative to when at least one source is neutral or friendly. In other words, if presented only with news items from undesirable sources, respondents would rather abstain than select a news item from an unfriendly source.

SMIV.4 Who Abstains?

We also assess whether the same demographic characteristics related to abstention in our F-M replication are related to abstention in our Mummolo replication. We use the demographic information we collected before the conjoint tasks and again construct an ordinary least squares model to regress a binary indicator for whether the respondent abstained in a given task (1) or not (0) on each of these characteristics except age and gender.[6] The estimates from this analysis are presented in Table SM.10. Some of the demographic characteristics generally thought to be related to abstention behave as expected; for instance, as respondents move up one point on the six-point income scale, they become one percentage point less likely to abstain in a given task. However, other demographic characteristics are related to abstention in our replication in ways that differ from how they are commonly associated with abstention in other settings; for instance, respondents identifying as Hispanic or Black are 10 percentage points less likely to abstain.

Overall, these results reflect those from our analogous analysis of our F-M replication

[6] We omit age and gender from this analysis they are used to code the “relevancy” attribute.

Table SM.10: Demographic correlates of abstention in Mummolo replication

	Model 1
Intercept	0.31^{*}
	(0.02)
Education	-0.01
	(0.01)
Asian	-0.03^{*}
	(0.02)
Black	-0.10^{*}
	(0.02)
Other Non-White Race	0.01
	(0.03)
Hispanic	-0.10^{*}
	(0.02)
Income	-0.01^{*}
	(0.00)
Ideology	0.00
	(0.00)
N	7958

Notes: Statistical significance is indicated as $^{*}p < 0.05$.

(see Section SMIII.4): individual-level characteristics are indicative of abstention such that the estimates associated with a conjoint experiment may differ systematically depending on whether an abstention option is present and the associations between specific demographic characteristics and abstention in conjoint experiments may not perfectly mirror the relationships observed in other survey settings.

SMIV.5 Switching Experimental Arms

After completing the twelve tasks in their assigned experimental arm, respondents were asked to complete a thirteenth, final task in which the outcome measure reflected that used in the opposite experimental arm (e.g., respondents in the forced-choice arm were now given the opportunity to choose one of the news items or to abstain). For each respondent, we randomly selected one of the twelve tasks they already completed with their assigned outcome measure and presented them with the same profiles as in that task but with the alternative outcome measure.

Of the 706 respondents in the forced-choice arm who completed all thirteen tasks (the twelve forced-choice tasks and the seventh abstention option task), 239 (33.9%) chose to abstain rather than choosing the candidate they originally selected when forced to do so. This descriptive evidence suggests that a fairly large proportion of respondents will choose news items if forced to do so even though they would ultimately prefer not to make a choice.

References

Arias, Sabrina B and Christopher W Blair. 2022. "Changing Tides: Public Attitudes on Climate Migration." *Journal of Politics* 84(1):560–567.

Auerbach, Adam Michael and Tariq Thachil. 2018. "How Clients Select Brokers: Competition and Choice in India's Slums." *American Political Science Review* 112(4):775–791.

Auerbach, Adam Michael and Tariq Thachil. 2020. "Cultivating Clients: Reputation, Responsiveness, and Ethnic Indifference in India's Slums." *American Journal of Political Science* 64(3):471–487.

Bakker, Bert N, Gijs Schumacher and Matthijs Rooduijn. 2021. "The Populist Appeal: Personality and Antiestablishment Communication." *Journal of Politics* 83(2):589–601.

Ballard-Rosa, Cameron, Lucy Martin and Kenneth Scheve. 2017. "The Structure of American Income Tax Policy Preferences." *Journal of Politics* 79(1):1–16.

Bansak, Kirk, Michael M Bechtel and Yotam Margalit. 2021. "Why Austerity? The Mass Politics of a Contested Policy." *American Political Science Review* 115(2):486–505.

Barnett, Carolyn, Amaney A. Jamal and Steve L Monroe. 2021. "Earned Income and Women's Segmented Empowerment: Experimental Evidence from Jordan." *American Journal of Political Science* 65(4):954–970.

Berinsky, Adam J. 2008. Survey Non-Response. In *The SAGE Handbook of Public Opinion Research*, ed. Wolfgang Donsbach and Michael W. Traugott. Sage pp. 309–321.

Blumenau, Jack and Benjamin E Lauderdale. 2022. "The Variable Persuasiveness of Political Rhetoric." *American Journal of Political Science* .

Campbell, Rosie, Philip Cowley, Nick Vivyan and Markus Wagner. 2019. "Why Friends and Neighbors? Explaining the Electoral Appeal of Local Roots." *Journal of Politics* 81(3):937–951.

Carnes, Nicholas and Noam Lupu. 2016. "Do Voters Dislike Working-Class Candidates? Voter Biases and the Descriptive Underrepresentation of the Working Class." *American Political Science Review* 110(4):832–844.

Costa, Mia. 2021. "Ideology, Not Affect: What Americans Want from Political Representation." *American Journal of Political Science* 65(2):342–358.

Dill, Janina and Livia I Schubiger. 2021. "Attitudes Toward the Use of Force: Instrumental Imperatives, Moral Principles, and International Law." *American Journal of Political Science* 65(3):612–633.

Doherty, David, Conor M Dowling and Michael G Miller. 2019. "Do Local Party Chairs Think Women and Minority Candidates Can Win? Evidence from a Conjoint Experiment." *Journal of Politics* 81(4):1282–1297.

Eggers, Andrew C, Nick Vivyan and Markus Wagner. 2018. "Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?" *Journal of Politics* 80(1):321–326.

Eshima, Shusei and Daniel M Smith. 2022. "Just a Number? Voter Evaluations of Age in Candidate-Choice Experiments." *Journal of Politics* .

Frederiksen, Kristian Vrede Skaaning. 2022. "Does Competence Make Citizens Tolerate Undemocratic Behavior?" *American Political Science Review* .

Funck, Amy S and Katherine T McCabe. 2022. "Partisanship, Information, and the Conditional Effects of Scandal on Voting Decisions." *Political Behavior* 44:1389–1409.

Ghosn, Faten, Tiffany S Chu, Miranda Simon, Alex Braithwaite, Michael Frith and Joanna Jandali. 2021. "The Journey Home: Violence, Anchoring, and Refugee Decisions to Return." *American Political Science Review* 115(3):982–998.

Hainmueller, Jens and Daniel J Hopkins. 2015. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes Toward Immigrants." *American Journal of Political Science* 59(3):529–548.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Hankinson, Michael. 2018. "When Do Renters Behave Like Homeowners? High Rent, Price Anxiety, and NIMBYism." *American Political Science Review* 112(3):473–493.

Hanretty, Chris, Benjamin E Lauderdale and Nick Vivyan. 2020. "A Choice-Based Measure of Issue Importance in the Electorate." *American Journal of Political Science* 64(3):519–535.

Häusermann, Silja, Thomas Kurer and Denise Traber. 2019. "The politics of trade-offs: Studying the dynamics of welfare state reform with conjoint experiments." *Comparative political studies* 52(7):1059–1095.

Henderson, John A, Geoffrey Sheagley, Stephen N Goggin, Logan Dancey and Alexander G Theodoridis. 2022. "Primary Divisions: How Voters Evaluate Policy and Group Differences in Intraparty Contests." *Journal of Politics* .

Johns, Robert and Ann-Kristin Kölln. 2020. "Moderation and Competence: How a Party's Ideological Position Shapes its Valence Reputation." *American Journal of Political Science* 64(3):649–663.

Kennedy, Ryan P, Philip D Waggoner and Matthew M Ward. 2022. "Trust in Public Policy Algorithms." *Journal of Politics* 84(2):1132–1148.

Levy, Gabriella. 2022. "Evaluations of Violence at the Polls: Civilian Victimization and Support for Perpetrators after War." *Journal of Politics* 84(2):783–797.

Liu, Guoer and Yuki Shiraito. 2023. "Multiple Hypothesis Testing in Conjoint Analysis." *Political Analysis* .

Magni, Gabriele. 2022. "Boundaries of Solidarity: Immigrants, Economic Contributions, and Welfare Attitudes." *American Journal of Political Science* .

Magni, Gabriele and Andrew Reynolds. 2021. "Voter Preferences and the Political Underrepresentation of Minority Groups: Lesbian, Gay, and Transgender Candidates in Advanced Democracies." *Journal of Politics* 83(4):1199–1215.

Mummolo, Jonathan. 2016. "News from the Other Side: How Topic Relevance Limits the Prevalence of Partisan Selective Exposure." *Journal of Politics* 78(3):763–773.

Mummolo, Jonathan and Clayton Nall. 2017. "Why Partisans Do Not Sort: The Constraints on Political Segregation." *Journal of Politics* 79(1):45–59.

Nelson, Michael J and Christopher Witko. 2021. "The Economic Costs of Democratic Backsliding? Backsliding and State Location Preferences of US Job-Seekers.".

Ono, Yoshikuni and Michael A Zilis. 2022. "Ascriptive Characteristics and Perceptions of Impropriety in the Rule of Law: Race, Gender, and Public Assessments of Whether Judges Can Be Impartial." *American Journal of Political Science* 66(1):43–58.

Peterson, Erik. 2017. "The Role of the Information Environment in Partisan Voting." *Journal of Politics* 79(4):1191–1204.

Peterson, Erik and Gabor Simonovits. 2018. "The Electoral Consequences of Issue Frames." *Journal of Politics* 80(4):1283–1296.

Poertner, Mathias. 2021. "The Organizational Voter: Support for New Parties in Young Democracies." *American Journal of Political Science* 65(3):634–651.

Rodon, Toni and Marc Sanjaume-Calvet. 2020. "How Fair Is It? An Experimental Study of Perceived Fairness of Distributive Policies." *Journal of Politics* 82(1):384–391.

Rosenstone, Steven J and John M Hansen. 2003. *Mobilization, Participation, and Democracy in America.* Pearson Education.

Schneider, Christina J. 2020. "Public Commitments as Signals of Responsiveness in the European Union." *Journal of Politics* 82(1):329–344.

Spater, Jeremy. 2022. "Exposure and Preferences: Evidence from Indian Slums." *American Journal of Political Science* 66(2):302–317.

Teele, Dawn Langan, Joshua Kalla and Frances Rosenbluth. 2018. "The ties that double bind: social roles and women's underrepresentation in politics." *American Political Science Review* 112(3):525–541.

Tellez, Juan Fernando. 2021. "Peace Without Impunity: Worldview in the Settlement of Civil Wars." *Journal of Politics* 83(4):1322–1336.

Tsai, Lily L, Minh Trinh and Shiyao Liu. 2022. "What Makes Anticorruption Punishment Popular? Individual-Level Evidence from China." *Journal of Politics* 84(1):602–606.

Ward, Dalston G. 2019. "Public Attitudes Toward Young Immigrant Men." *American Political Science Review* 113(1):264–269.

Weaver, Julie Anne. 2021. "Electoral Dis-Connection: The Limits of Reelection in Contexts of Weak Accountability." *Journal of Politics* 83(4):1462–1477.