The Limitations of Using Forced Choice in Electoral Conjoint Experiments *

Giancarlo Visconti[†] Yang Yang[‡]

Abstract

Electoral conjoint experiments typically require respondents to choose between two candidates, assuming full turnout and strict preference orderings, even though real-world voters may abstain or cast protest votes. We show that excluding these options is not trivial. In our preregistered experiment, more than half of respondents abstained or cast protest votes at least once when such options were available. While most AMCE estimates remain stable, a subset shifts in direction, magnitude, or significance—changes that are unpredictable ex ante and can mislead inferences about voter preferences. Marginal mean estimates diverge even sharply across designs. These shifts are systematic, shaped by respondent-level traits (e.g., turnout history and propensity) and task-level features such as rating differentials and contest appeal. Our findings show that forced-choice designs impose strong behavioral assumptions and risk obscuring meaningful variation in electoral behavior. Providing explicit opt-out options enhances validity and better captures the full range of voter preferences.

Keywords: Conjoint Experiment, Voter Preferences, Forced Choice, Bias

*Authors are listed alphabetically. We thank Alex Coppock, Ariadna Chuaqui, Bernardo Lara, Bumba Mukherjee, Carmen Le Foulon, Cory McCartan, Jacob Turner, Justin Esarey, Pablo Pinto, Santiago López-Cariboni, Santiago Olivella, Tara Slough, and seminar participants at Penn State C-SoDA, the NYU Rebecca B. Morton Conference on Experimental Political Science, University of Maryland, Latin American PolMeth 2024, and APSA 2024 for their helpful comments and suggestions. The survey was implemented under Penn State University IRB Protocol STUDY0002477. This project was funded by a C-SoDA Accelerator Award. The design was registered at OSF before data collection. All errors are our own.

[†]Assistant Professor, Department of Government & Politics, University of Maryland, College Park; gvis@umd.edu.

[‡]Ph.D. Candidate, Department of Political Science and Center for Social Data Analytics, Pennsylvania State University;yky5272@psu.edu

Understanding how voters choose between candidates with varying characteristics is a core question in electoral research. Conjoint analysis has become popular for examining voter decisions. Whereas real world elections often allow for abstention or protest voting, most electoral conjoint studies employ a forced-choice design, requiring participants to choose one candidate over the other.[1] This mismatch between experimental design and real-world electoral behavior raises fundamental questions about what conjoint experiments actually reveal about voter preferences. Should conjoint tasks offer opt-out choices—such as abstention or "None of these"—to capture the full decision space voters face? If introduced, would respondents who were previously "forced" to choose instead select these alternatives? And critically, are such shifts merely random noise, or do they reflect systematic and meaningful patterns in voter preferences?

This study addresses these concerns by investigating whether, and under what conditions, the forced-choice design in electoral conjoint experiments can constrain the expression of voter preferences.[2] We demonstrate that employing a forced-choice design is not merely a design nuance but one with theoretical and empirical consequences. To motivate this inquiry, we first revisit two assumptions embedded in the forced-choice design: first, that each respondent must choose one preferred profile within each choice task (Hainmueller et al., 2014); and second, that each respondent holds a strict preference ordering over all unique candidates, implying no indifferent preferences

[1]We surveyed articles using conjoint designs to study voter choices between hypothetical candidates. Of the 72 articles identified, only four did not employ a forced-choice design. A detailed list can be found in Appendix A.

[2] One might argue that abstention in real-world elections primarily reflects turnout costs, and because conjoint experiments eliminate such costs, there is no need to include an abstention option. However, if turnout costs were the sole determinant of abstention and were truly removed in conjoint settings, we should observe that no respondents choose to abstain, even when the option is available. Yet, if respondents still opt to abstain and do so systematically rather than at random, this suggests that abstention is not merely a function of turnout costs but instead reflects deeper attitudinal or evaluative considerations that conjoint designs should capture.

(Abramson et al., 2022; Bansak et al., 2023).[3] Using existing conjoint data from published articles, we demonstrate that these assumptions can be violated. First, a substantial proportion of respondents recruited for electoral conjoint experiments report prior abstention despite being eligible. Second, nearly half assign identical ratings to both candidates, suggesting they may choose neither or abstain if given the option. Such patterns suggest that forced-choice conjoint experiments may create artificial trade-offs, distorting inferences about voter preferences.

Drawing on electoral scholarship (for a review, see Alvarez et al., 2018), we argue that the decisions to vote or abstain and to choose among listed candidates or cast a protest vote constitute distinct and meaningful dimensions of voter preferences. Ignoring these options in forced-choice conjoint designs generates two forms of design-induced bias: misclassification bias and external validity bias. Excluding protest-vote options creates misclassification bias, as respondents who would reject both candidates are compelled to choose one, thereby distorting their expressed preferences. Omitting abstention introduces external validity bias by forcing participation from individuals who would not vote in real elections, thus misaligning the experimental sample with the target electorate.

To assess the magnitude of these biases and the patterns of opting out, we conduct a preregistered randomized experiment that assigns participants to either a standard forced-choice design or an unforced-choice design. Each design embeds two replicated electoral conjoint studies: one on presidential candidates from Hainmueller et al. (2014) and another on congressional candidates from Peterson (2017). We also include pre-conjoint measures capturing respondents' past voting behavior, voting propensity, and demographic characteristics. This unique design enables us to not only compare how introducing abstention and protest-vote options affects preference estimates relative to the conventional forced-choice design, but also to examine which profile-, task-, and respondent-level factors predict switching behavior in an unforced-choice setting.

[3]Together, these assumptions imply the Independence of Irrelevant Alternatives (IIA): additional options (e.g., abstention or protest voting) are presumed irrelevant, and their inclusion will not alter the estimated probabilities of choosing between existing candidates.

First, across both replicated studies, roughly half of respondents used the abstention or protest-vote options at least once under the unforced-choice design, and these behavioral shifts produced measurable differences in conjoint estimates. Most AMCE estimates remained stable, but a subset showed modest yet substantively meaningful changes—differences in magnitude, direction, or statistical significance in Study 1, and smaller, nonsignificant shifts in Study 2. However, the impact of design is much more pronounced for marginal means (MMs), which capture absolute support rather than relative preferences, given that aggregate levels of expressed support differ sharply between forced- and unforced-choice settings. To ensure the design-induced differences are not driven by selection into abstention, we apply Inverse Probability Weighting (IPW) and two-step Heckman selection models. In all cases, the substantive conclusions remain unchanged.

Second, respondents' use of opt-out options is highly systematic rather than random or incidental. At the individual level, past voting behavior, prospective turnout intentions, and age strongly predict switching, indicating that respondents draw on the same underlying dispositions that shape participation in real elections. At the task level, switching is most likely when the two profiles are perceived as indistinguishable in appeal—an unmistakable signal of indifference. Conversely, when both profiles are evaluated negatively (rated below the midpoint), respondents are substantially more likely to switch, consistent with disaffection or protest behavior. These patterns closely parallel established theories of turnout and protest voting. Crucially, once respondent- and task-level factors are taken into account, most candidate attributes show limited association with switching—particularly abstention. This suggests that opt-out behaviors in conjoint tasks are shaped more by respondents' broader political dispositions and their overall evaluations of the contest than by any single attribute manipulation. Far from being noise, these expressions reflect structured behavioral responses that forced-choice designs are unable to capture.

Recent work highlights that conjoint estimates can be distorted by various sources of error. Clayton et al. (2025) and Kane and Costa (2024) emphasize respondent-driven errors from unreliability or inattentiveness, whereas we focus on design-driven biases that persist even among attentive respondents and under unbiased estimators. Our study is closely related to Miller and

Ziegler (2024), who assess how including an abstention option affects conjoint estimates, and to Hainmueller et al. (2015), who show that non-forced-choice designs better approximate real-world behavior, but leave the behavioral mechanisms unexplained.[4] We identify the behavioral mechanisms behind these shifts and show that opting out is neither rare nor random; it is a structured electoral behavior that voters deploy in predictable ways, with direct consequences for how conjoint experiments capture voter preferences.

This paper contributes to the growing literature on conjoint analysis in political methodology and to the study of elections and voting behavior. While methodological advances have refined estimands (e.g., Leeper et al., 2020; Abramson et al., 2022; Zhirkov, 2022; Ganter, 2023), clarified interpretation (Abramson et al., 2022; Bansak et al., 2023), and improved validity and response quality (Bansak et al., 2018; Horiuchi et al., 2021; Clayton et al., 2025; Hainmueller et al., 2015; de la Cuesta et al., 2022; Fu and Li, 2025), our study revisits a more fundamental concern: how outcomes are defined and what behavioral assumptions they entail. In electoral studies, forced-choice designs compel respondents to choose one candidate per task, implicitly assuming full turnout and strict preference orderings (Hainmueller et al., 2014; Abramson et al., 2022; Bansak et al., 2023). When violated, these assumptions can distort inferences about voter preferences.[5] Although recent studies adopt unforced-choice designs (e.g., Eggers et al., 2018), empirical evidence on their implications remains limited. Our study demonstrates that conjoint experiments can solicit not only preference rankings but also meaningful forms of non-choice—an important dimension of

[4]Miller and Ziegler (2024) find that abstention increases when a candidate scandal is present, especially involving a copartisan, and that information complexity reduces abstention only in such cases. By contrast, we examine more general task-level features—rating differentials and contest appeal—and respondent-level characteristics that extend beyond specific attribute content.

[5]Even when the substantive quantity of interest seeks to capture preferences rather than expected vote shares, the forced-choice format still assumes that one of the two candidates is preferred. Yet indifference or preferring none of the listed profiles is itself a meaningful preference that forced-choice designs cannot capture.

political preferences relevant to elections, immigration, welfare programs, and other topics where individuals may prefer "none of the above."

Our findings also carry important implications for substantive studies of electoral behavior. It supports the idea that nonparticipation and protest voting are not residual behaviors but meaningful political acts shaped by both contest appeal and individual traits. By showing how respondents navigate political choices when none of the available options are appealing, our results speak to broader theories of political alienation and the conditions under which citizens withdraw from electoral choice. Allowing abstention and protest options reveals patterned forms of indifference, dissatisfaction, and disengagement that forced-choice designs mask.

The AMCE as a Causal Quantity of Interest

The AMCE is a central estimand in conjoint analysis. When estimated using a design-based approach grounded purely in randomization, it can be interpreted as the average causal effect of a candidate attribute on the expected vote share in an electoral conjoint study. This interpretation is formalized and justified in both Abramson et al. (2022) and Bansak et al. (2023). Formally, consider a paired-profile, forced-choice conjoint experiment in which each respondent $i \in \{1,\ldots,N\}$ completes K tasks, choosing between two profiles. Each profile is defined by a vector of L attributes, each with discrete levels. For illustration, suppose profiles are defined by three binary attributes A, B, and C, so that $[abc]$ denotes a candidate profile with $A = a$, $B = b$, and $C = c$.

Let $Y_i([abc],[a'b'c']) \in \{0,1\}$ denote respondent i's potential outcome in a task comparing profiles $[abc]$ and $[a'b'c']$. The outcome equals 1 if the respondent prefers $[abc]$ over $[a'b'c']$, and 0 otherwise. The AMCE for attribute A is then defined as:

$$\text{AMCE}_A = \mathbb{E}\left[Y_i([1BC],[A'B'C']) - Y_i([0BC],[A'B'C'])\right], \tag{1}$$

where the expectation is taken over the joint distribution of all other attributes (i.e., B, C, A', B', C') as defined by the experimental randomization. This estimand captures the average effect of changing A from 0 to 1 on the probability that a candidate wins a hypothetical head-to-head election, averaging over the distribution of other features and across respondents.

This design-based estimand is especially meaningful in electoral research. Treating each profile comparison as a stylized election, the AMCE captures the average change in a candidate's vote share resulting from altering a particular attribute A. As Abramson et al. (2022) demonstrate, this quantity can be understood as the aggregation of many simulated pairwise elections—each reflecting a respondent's individual comparison—such that the AMCE summarizes the overall pattern of how attribute changes would influence outcomes across these simulated contests, similar in spirit to how a Borda count aggregates individual rankings into a collective preference. Bansak et al. (2023) generalize this by showing that, under random assignment and sincere voting, the AMCE identifies the marginal effect of an attribute on choice probabilities, corresponding to the expected vote share.

To formalize this connection, Bansak et al. (2023) describe how AMCE depends on two target distributions: the attribute distribution $\mathscr{A}$ and the voter preference distribution $\mathscr{V}$. Here, $\mathscr{A}$ represents how candidate profiles are drawn—i.e., the probability measure over feature combinations—while $\mathscr{V}$ captures how respondents (with specific preference orderings) are sampled in the experiment. When a conjoint experiment randomizes profiles according to $\mathscr{A}$ and recruits respondents according to $\mathscr{V}$, the AMCE reflects the expected change in vote share that results from switching a candidate's attributes (e.g., gender or policy stance), holding other attributes constant.

While recent research has explored how $\mathscr{A}$ affects external validity (de la Cuesta et al., 2022), less attention has been paid to $\mathscr{V}$. Forced-choice conjoint designs assume full participation and strict orderings by requiring respondents to choose between Candidate A and B in every task (Hainmueller et al., 2014; Abramson et al., 2022; Bansak et al., 2023).[6] When applied to samples

[6]Countries and subnational units like India, Ukraine, Bulgaria, Thailand, Colombia, and the U.S. state of Nevada offer the NOTA option. The French equivalent, 'vote blanc,' and 'voto blanco' in some Latin American countries serve a similar purpose. See Alvarez et al. (2018) and Plescia et al. (2023) for more details.

meant to represent a real electorate, these designs compel real-world abstainers and indifferent voters to state choices they do not genuinely hold, artificially inflating support and weakening external validity. These concerns apply even when the estimand is preference rather than expected vote share. The forced-choice format still presumes that one of the two candidates is preferred, yet indifference—or preferring none of the listed profiles—is itself a meaningful preference that forced-choice designs cannot accommodate. Ignoring such options risks systematically misrepresenting how citizens evaluate political alternatives.

Not All Respondents Turn Out to Vote

Not all eligible individuals vote, and abstention is not random (Bernhagen and Marsh, 2007). Research shows that non-voters systematically differ from voters in age, education, mobilization, party ID, political interest, and knowledge (e.g., Adams and Merrill, 2003; Bernhagen and Marsh, 2007; Battaglini et al., 2008; Koch et al., 2023; Smets and van Ham, 2013). If all eligible citizens turned out, electoral outcomes could shift—highlighting the relevance of counterfactual preference aggregation (Bernhagen and Marsh, 2007; Kawai et al., 2021). Nonetheless, forced-choice conjoint experiments assume full turnout, requiring respondents to select a candidate in each task (Hainmueller et al., 2014), even if they might abstain in real elections. By incorporating non-voters without accounting for their distinct preferences, these designs risk biasing vote choice estimates.

Do real-world non-voters participate in conjoint studies? While most designs do not capture this, Horiuchi et al. (2020) include a post-conjoint question on voting history.[7] Of 2,200 respondents, 600 reported always voting, while over 370 almost never did. Re-estimating AMCEs by group reveals notable differences: habitual non-voters placed greater weight on prior government experience than consistent voters (Figure 1). These results suggest that non-voter preferences differ systematically, and their inclusion without adjustment may distort experimental findings.

[7] Q7.8: "How often have you participated in voting since you got the right to vote?" Responses: (1) All elections, (2) Most, (3) Several, (4) Almost never. Translated from Japanese.

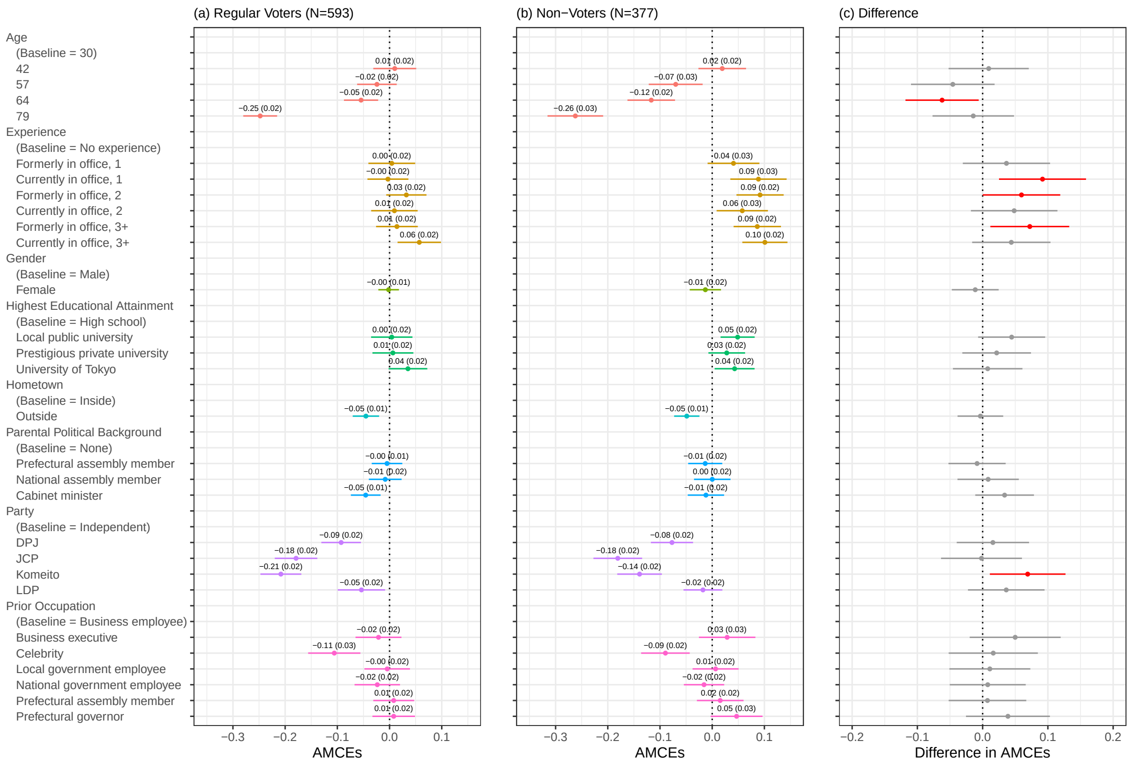

Figure 1: Replication of the Horiuchi et al. (2020) experiment using subsamples of regular and non-voters. N denotes the number of respondents in each subsample. The rightmost panel displays differences in AMCEs between regular and non-voters for each attribute level, with horizontal bars representing 95% confidence intervals, adjusted for clustering at the respondent level. Red dots with red bars indicate significance at the 5% level, red triangles with gray bars indicate significance at the 10% level, and gray dots with gray bars indicate non-significant results at conventional levels.

Respondents Hold Indifferent Preferences

Protest voting—such as casting a blank or null ballot—is a deliberate act by politically engaged individuals who reject the available options while remaining committed to participation (Cohen, 2018). In contrast, abstention is more often linked to disengagement, low political interest, or a perceived lack of difference between candidates (Blais and Daoust, 2020). They reflect different responses to dissatisfaction and challenge the assumption that all voters/respondents hold strict preferences (Abramson et al., 2022; Bansak et al., 2023).

While binary outcomes conceal such sentiments, rating-based responses—intended to capture preference intensity (Hainmueller et al., 2014)—provide a useful proxy: when respondents assign identical scores to both profiles, they likely felt indifferent but were forced to choose. In Hainmueller et al. (2014) and Kirkland and Coppock (2018), more than 40% of respondents gave identical ratings in at least one task (Figure 2). This likely underestimates true indifference, as some may have adjusted ratings to align with their forced choices.[8] In a realistic setting, many might have abstained or protested rather than made an artificial trade-off.

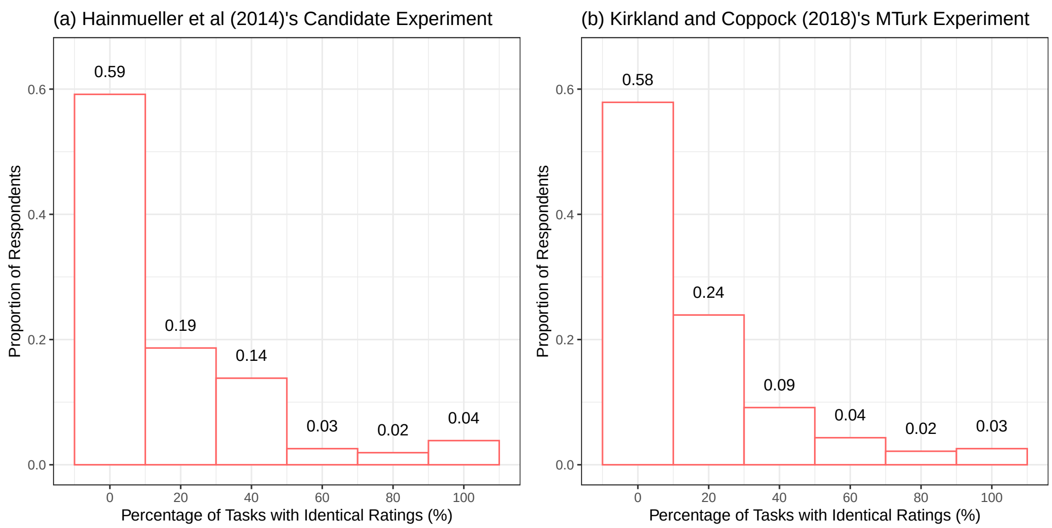

Figure 2: Proportion of Respondents Assigning Identical Ratings to Tasks in Hainmueller et al. (2014)'s and Kirkland and Coppock (2018)'s Experiments

Empirical Implications of Opt-Out Options in Conjoint Studies

Electoral research distinguishes between casting a blank or null vote and abstaining, as each reflects distinct motivations and individual characteristics (Alvarez et al., 2018; Cohen, 2018). We use stylized examples to show how this conceptual distinction translates into empirically meaningful differences in conjoint analysis through outcome coding. Consider an electorate of ten voters in a conjoint task simulating an election that allows for abstention and protest votes but imposes a

[8] The number of tied pairs might be further underestimated because, in both Hainmueller et al. (2014) and Kirkland and Coppock (2018), the rating questions were not mandatory. Many respondents skipped one or both rating items in a task, and these missing responses were excluded from the count of identical ratings.

forced-choice format. In the first example (Table 1), Voter 1 intends to cast a protest vote (e.g., a blank or invalid ballot). [9]

Because protest votes count toward the total votes cast but not toward any candidate (vote for candidate), both profiles are coded as 0. Under the forced-choice format, they are required to select a candidate, resulting in a recorded vote of 1 for Candidate t_b. This misrepresents their true intention and introduces misclassification bias. The equations at the bottom of Table 1 illustrate how this leads to different estimates of vote share in a stylized election.

Table 1: Toy Example 1.

	Survey			Reality		
	Turnout=100% Valid Votes=100%			Turnout=100% Valid Votes=90%		
Eligible Voter i	$Y_i(t_a)$	$Y_i(t_b)$	π_i	$\tilde{Y}_i(t_a)$	$\tilde{Y}_i(t_b)$	$\tilde{\pi}_i$
1	0	1	-1	0	0	0
2	1	0	1	1	0	1
3	1	0	1	1	0	1
4	0	1	-1	0	1	-1
5	0	1	-1	0	1	-1
6	1	0	1	1	0	1
7	0	1	-1	0	1	-1
8	1	0	1	1	0	1
9	1	0	1	1	0	1
10	1	0	1	1	0	1
Counts	6	4	2	6	3	3

$$\begin{aligned} E[Y_i(t_a) - Y_i(t_b)] &= E[Y_i(t_a)] - E[Y_i(t_b)] \\ &= \frac{6}{10} - \frac{4}{10} = \frac{1}{5} \\ E[\tilde{Y}_i(t_a) - \tilde{Y}_i(t_b)] &= E[\tilde{Y}_i(t_a)] - E[\tilde{Y}_i(t_b)] \\ &= \frac{6}{10} - \frac{3}{10} = \frac{3}{10} \end{aligned} \tag{2}$$

Similarly, when abstention is substantively meaningful, forced-choice designs misrepresent not only preferences but also participation. In the second toy example (Table 2), Voter 1 would abstain in a real election, contributing neither to turnout nor to any candidate's vote share. In a forced-choice conjoint, however, this voter is compelled to choose, and their response is recorded as a vote for Candidate t_b. This artificial participation introduces external validity bias by reshaping the study population to include individuals who, under real electoral conditions, would not vote at all.

[9] Recall that vote share is calculated as Vote share $= \frac{\text{Number of votes for candidate}}{\text{Total votes cast}} \times 100$.

Table 2: Toy Example 2.

	Survey			Reality		
	Turnout=100% Valid Votes=100%			Turnout=90% Valid Votes=100%		
Eligible Voter i	$Y_i(t_a)$	$Y_i(t_b)$	π_i	$\tilde{Y}_i(t_a)$	$\tilde{Y}_i(t_b)$	$\tilde{\pi}_i$
1	0	1	-1	-	-	-
2	1	0	1	1	0	1
3	1	0	1	1	0	1
4	0	1	-1	0	1	-1
5	0	1	-1	0	1	-1
6	1	0	1	1	0	1
7	0	1	-1	0	1	-1
8	1	0	1	1	0	1
9	1	0	1	1	0	1
10	1	0	1	1	0	1
Counts	6	4	2	6	3	3

$$\begin{aligned} E[Y_i(t_a) - Y_i(t_b)] &= E[Y_i(t_a)] - E[Y_i(t_b)] \\ &= \frac{6}{10} - \frac{4}{10} = \frac{1}{5} \\ E[\tilde{Y}_i(t_a) - \tilde{Y}_i(t_b)] &= E[\tilde{Y}_i(t_a)] - E[\tilde{Y}_i(t_b)] \\ &= \frac{6}{9} - \frac{3}{9} = \frac{1}{3} \end{aligned} \tag{3}$$

As shown above, protest voting and abstention carry distinct electoral implications, and this distinction should guide their coding in conjoint experiments when the substantive quantity of interest is expected vote share. Protest votes—blank, null, or spoiled ballots—count toward turnout but not toward any candidate; we therefore code them as 0 for both candidates. Abstention reflects non-participation and is excluded from both the numerator and denominator of vote-share calculations, so we code it as missing.[10]

This distinction clarifies two sources of design-induced bias in forced-choice conjoint designs. Misclassification bias occurs when respondents who would cast a protest vote are instead forced to select a candidate. External validity bias arises when respondents who would abstain are compelled to "participate," thereby altering the effective study electorate. Figure 3 illustrates how these biases distort inference.

[10]By contrast, Miller and Ziegler (2024) code abstention as zero for all profiles without distinguishing it from protest voting. We differentiate between these behaviors to preserve their distinct electoral and behavioral meanings.

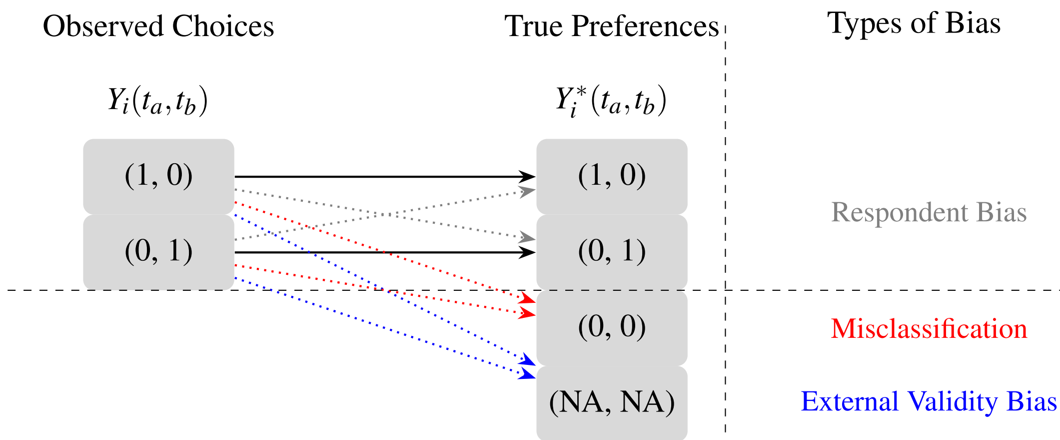

Figure 3: Types of Bias in Forced-Choice Electoral Conjoint Experiments. This figure illustrates the biases that arise in forced-choice electoral conjoint designs when the observed choice $Y_i(t_a, t_b)$ diverges from the respondent's true preference $Y_i^*(t_a, t_b)$. Because respondents must select one of the two profiles, true behaviors such as abstention (NA, NA) or protest voting (0,0) cannot be expressed and therefore remain unobserved, leading to systematic discrepancies between recorded and actual preferences.

Misclassification

Misclassification bias in conjoint experiments is a form of measurement error in choice-based outcomes. It has two sources: respondent-driven misclassification, arising from inconsistent or inattentive choices (Clayton et al., 2025; Kane and Costa, 2024), and design-driven misclassification, which occurs when respondents who genuinely prefer neither option are compelled to choose, generating incorrect outcome values. Although measurement error in dependent variables is often assumed to be innocuous, this applies only to classical error. Misclassification in binary outcomes is a non-classical measurement error and therefore systematically biases estimates (Meyer and Mittag, 2017; Shu and Yi, 2019).

In an electoral conjoint experiment, where the key outcome variable is binary, let the true binary choice-based outcome be Y^* , where $Y_i^* \in \{0, 1\}$ corresponds to each profile i in a paired electoral conjoint task. Each profile includes multiple attributes with varying levels. As introduced by Hainmueller et al. (2014) and Leeper et al. (2020), the primary quantity of interest is the profile-level AMCE, which is the difference in marginal means (MM) between two attribute levels. Profile-level MMs can be estimated nonparametrically using linear regression:

$$Y_i^* = \beta_0 + \sum_j \beta_j^{(k)} D_j^{(k)} + \varepsilon_i \tag{4}$$

where $\beta_j^{(k)}$ represents the true MM of attribute j's level k on the outcome; $D_j^{(k)}$ is an indicator variable that equals 1 if profile i, randomly generated by design, has attribute j at level k , and 0 for all other levels of j; and ε_i is the error term clustered at the respondent level, assumed to have a mean of 0 and be independent of $D_j^{(k)}$ due to randomization.

Now, let's consider the scenario where the binary outcome Y^* (the true voting choice) is misclassified as the observed binary outcome Y. Let the observed outcome be denoted as Y (the mismeasured voting choice), subject to both respondent- and design-induced measurement error. We define the probabilities of false negatives and false positives in misclassification, conditional on the true response and the attribute level $D_j^{(k)}$, as follows:[11]

$$P(Y_i = 0 | Y_i^* = 1, D_j^{(k)}) = p_{10} \tag{5}$$

$$P(Y_i = 1 | Y_i^* = 0, D_j^{(k)}) = p_{01} \tag{6}$$

The probabilities of no misclassification are:

$$P(Y_i = 1 | Y_i^* = 1, D_j^{(k)}) = 1 - p_{10} \tag{7}$$

$$P(Y_i = 0 | Y_i^* = 0, D_j^{(k)}) = 1 - p_{01} \tag{8}$$

With misclassification, according to the potential outcome framework in causal inference, the

[11] To clarify, respondent-driven misclassification—such as intra-respondent unreliability Clayton et al. (2025) or general inattentiveness (Kane and Costa, 2024)—can generate both types of error: false negatives (p_{10}) and false positives (p_{01}). In contrast, design-driven misclassification, which arises when respondents are forced to choose between two candidates despite preferring neither, primarily induces false positives ($p_{01} > 0$), as those who would otherwise abstain or cast a protest vote are compelled to select a candidate.

observed outcomes Y_i given the attribute level $D_j^{(k)}$ are realized as:

$$\begin{aligned} MM_j^k = \mathbb{E}(Y_i|D_j^{(k)}=1) = P(Y_i=1|D_j^{(k)}=1) \\ &= P(Y_i=1, Y_i^*=1|D_j^{(k)}=1) + P(Y_i=1, Y_i^*=0|D_j^{(k)}=1) \\ &= (1-p_{10})P(Y_i^*=1|D_j^{(k)}=1) + p_{01}P(Y_i^*=0|D_j^{(k)}=1) \\ &= (1-p_{10})(\beta_0+\beta_j^{(k)}) + p_{01}(1-\beta_0-\beta_j^{(k)}) \end{aligned} \tag{9}$$

For the reference level $D_j^{(k')}$, the outcomes are derived similarly:

$$\begin{aligned} MM_j^{k'} = \mathbb{E}(Y_i|D_j^{(k')}=1) = P(Y_i=1|D_j^{(k')}=1) \\ &= P(Y_i=1, Y_i^*=1|D_j^{(k')}=1) + P(Y_i=1, Y_i^*=0|D_j^{(k')}=1) \\ &= (1-p_{10})(\beta_0+\beta_j^{(k')}) + p_{01}(1-\beta_0-\beta_j^{(k')}) \end{aligned} \tag{10}$$

As the AMCE is causally defined as the difference in the population probabilities of choosing the J profiles that would result if a respondent were shown the profiles with attribute j's k level as opposed to k' level, we can compute the difference as follows:

$$\begin{aligned} AMCE_j^{kk'} &= MM_j^k - MM_j^{k'} \\ &= \mathbb{E}(Y_i|D_j^{(k)}=1) - \mathbb{E}(Y_i|D_j^{(k')}=1) \\ &= P(Y_i=1|D_j^{(k)}=1) - P(Y_i=1|D_j^{(k')}=1) \\ &= (1-p_{10}-p_{01})(\beta_j^{(k)}-\beta_j^{(k')}) \end{aligned} \tag{11}$$

Thus, as shown in Equation (11), with misclassification, the estimated AMCE is biased by a factor of $(1-p_{10}-p_{01})$. The true AMCE, which would have been $\beta_j^{(k)}-\beta_j^{(k')}$, is scaled by this misclassification factor. If there is no misclassification, then the estimated AMCE is equal to the true AMCE. This implies that even in the absence of respondent-driven errors ($p_{10}=0$), forced-choice designs can still bias estimates downward, attenuating the estimated AMCE toward zero.[12] How-

[12]Although this bias is functionally similar to attenuation bias, that term is conventionally reserved for measurement error in explanatory variables (Wooldridge, 2010); to avoid confusion, we refer to this problem as design-induced *misclassification bias*.

ever, if both types of misclassification occur, the estimates will be further biased and may go in any direction.

External Validity Bias

External validity bias occurs when the study sample, shaped by enrollment processes and inclusion/exclusion criteria, differs from the target sample representing the broader population (Degtiar and Rose, 2023). These differences can lead to biases when generalizing findings from the study sample to the target population.

For clarity, we assume that the distribution of candidate attributes $D_j^{(k)}$ is identical across study and target populations, but turnout behavior differs. The target population P_T includes only voters who turn out, while the study population P_S includes all eligible voters. This mismatch introduces bias if individuals who would abstain in real elections are nonetheless forced to choose in the experiment.[13] In an electoral conjoint study, the target population, P_T , consists of all eligible voters who actually turn out, while the study population, P_S, includes all eligible voters, regardless of whether they turn out or not.

The true AMCE for the target population of voters who turned out is denoted as:

$$\text{AMCE}_j^{(kk')}(P_T) = \mathbb{E}_{P_T}[Y_i|D_j^{(k)} = 1] - \mathbb{E}_{P_T}[Y_i|D_j^{(k')} = 1] \tag{12}$$

Whereas the AMCE for the study population, which includes both those who may turn out and those who may not, is:

$$\text{AMCE}_j^{(kk')}(P_S) = \mathbb{E}_{P_S}[Y_i|D_j^{(k)} = 1] - \mathbb{E}_{P_S}[Y_i|D_j^{(k')} = 1] \tag{13}$$

[13] Although candidate attributes may influence abstention, our analysis shows that choosing abstention is primarily driven by respondent- and task-level factors rather than single profile attributes. Adjusting for this potential selection using IPW and Heckman models does not change the main results (see Appendix C.6).

We suggest that external validity bias arises when the AMCE calculated using the study population differs from the AMCE in the target population. Mathematically, the bias can be expressed as:

$$\text{External Validity Bias}_j^{(kk')} = \text{AMCE}_j^{(kk')}(P_S) - \text{AMCE}_j^{(kk')}(P_T) \tag{14}$$

Let's assume that the outcome Y_i in the conjoint experiment (e.g., vote choice) depends not only on the profile attributes $D_j^{(k)}$ but also on whether or not an individual turns out to vote in each pair of comparisons. Let:

- $T_i = 1$ denotes that the individual prefers to turn out to vote.
- $T_i = 0$ denotes that the individual prefers not to turn out to vote.

We assume that preferences (reflected in Y_i) differ between those who turn out and those who do not, but the distribution of profile attributes $D_j^{(k)}$ is the same between the study and target populations.

The study population, P_S, includes both voters $(T_i = 1)$ and non-voters $(T_i = 0)$. Therefore, the expected outcome in the study population is a weighted average of the outcomes for voters and non-voters. These weights are the probabilities $\Pr(T_i = 1|P_S)$ (the probability of being a voter in the study population) and $\Pr(T_i = 0|P_S)$ (the probability of being a non-voter).

$$\begin{aligned}
\text{AMCE}_j^{(kk')}(P_S) &= \left(\mathbb{E}_{P_S}[Y_i|D_j^{(k)} = 1] - \mathbb{E}_{P_S}[Y_i|D_j^{(k')} = 1]\right) \\
&= \Pr(T_i = 1|P_S) \cdot \left(\mathbb{E}[Y_i|D_j^{(k)} = 1, T_i = 1] - \mathbb{E}[Y_i|D_j^{(k')} = 1, T_i = 1]\right) \\
&+ \Pr(T_i = 0|P_S) \cdot \left(\mathbb{E}[Y_i|D_j^{(k)} = 1, T_i = 0] - \mathbb{E}[Y_i|D_j^{(k')} = 1, T_i = 0]\right)
\end{aligned} \tag{15}$$

The target population includes only voters $(T_i = 1)$. Therefore, the AMCE in the target population is based solely on the expected outcomes for voters:

$$\text{AMCE}_j^{(kk')}(P_T) = \mathbb{E}[Y_i|D_j^{(k)} = 1, T_i = 1] - \mathbb{E}[Y_i|D_j^{(k\prime)} = 1, T_i = 1] \tag{16}$$

We can now substitute Equation (14) with Equations (15) and (16). Notice that the voter terms

$\Pr(T_i = 1|P_S)$ can partially cancel out with the corresponding voter terms in the target population. The remaining bias is driven by the non-voter terms.

$$
\begin{aligned}
\text{External Validity Bias}_j^{(kk')} &= \text{AMCE}_j^{(kk')}(P_S) - \text{AMCE}_j^{(kk')}(P_T) \\
&= (\Pr(T_i = 1|P_S) - 1) \cdot \left(\mathbb{E}[Y_i|D_j^{(k)} = 1, T_i = 1] - \mathbb{E}[Y_i|D_j^{(k')} = 1, T_i = 1]\right) + \\
&\Pr(T_i = 0|P_S) \cdot \left(\mathbb{E}[Y_i|D_j^{(k)} = 1, T_i = 0] - \mathbb{E}[Y_i|D_j^{(k')} = 1, T_i = 0]\right)
\end{aligned}
\tag{17}
$$

Equation (17) shows that the external validity bias depends on two factors: 1) The proportion of non-voters in the study population, $\Pr(T_i = 0|P_S)$. 2) The mean difference in preferences (i.e., the expected outcomes) between voters and non-voters for each attribute level. In essence, if non-voters have different mean preferences compared to voters, this will introduce bias when attempting to generalize the AMCE from the study population to the target population. Intuitively, the more non-voters there are in the study population (sample), the larger the potential bias.

Experiment Design

To assess bias patterns in forced-choice conjoint designs, we conduct a preregistered randomized experiment embedding two conjoint studies. Respondents are randomly assigned to either a forced-choice design (control) or an unforced-choice design (treatment). In the forced-choice condition, we follow the exact design introduced by Hainmueller et al. (2014): respondents answer a binary choice question between two candidates, which they cannot skip, followed by two optional rating questions.[14] In the unforced-choice condition, we add two additional response options that explicitly capture non-choice behaviors—“If I only have these two candidates, I will cast

[14] Although Hainmueller et al. (2014) do not explicitly describe the binary choice question as a forced response and the rating questions as optional, their theoretical framework assumes that all respondents must make a choice and their implementation enforces this requirement in practice, as confirmed by our examination of their replication data. The same approach is adopted in other studies, such as Kirkland and Coppock (2018).

a blank/null vote"[15] and "If I only have these two candidates, I will abstain." These options are worded to distinguish between protest voting and abstention.[16] Unlike the forced-choice design, we make the rating items mandatory to encourage respondents to evaluate each candidate even when they prefer not to choose between them.[17]

Both designs rely on the same candidate profile generation but differ in how choice-based response options are presented. The experiment was fielded in June 2024 with a nationally representative U.S. sample of 2,704 respondents recruited via Cloud Research using quotas for age (18+), gender, ethnicity, and region. The survey begins with questions on voting history, turnout likelihood, demographics, and political attitudes. Respondents then complete eight paired evaluations of hypothetical presidential and congressional candidates, with the order of blocks randomized. Each

[15]We also considered phrasing protest behavior as "I would vote for a third-party candidate." But in the context of U.S. federal elections—where third-party options are far more familiar than blank ballots—this wording risked artificially inflating protest voting by drawing respondents away from the two major parties. To avoid overstating protest behavior, we instead used "I would cast a blank/null vote," which more clearly captures expressive disapproval while still requiring participation.

[16]We do not hide the abstention option because skipping a question may indicate inattentiveness or non-cooperation. Making abstention explicit ensures it reflects a conscious choice. One concern is that offering abstention and protest-vote options might prime behaviors not always visible on a real ballot. While reasonable, this concern is limited: respondents with genuine preferences should not be diverted by additional options, whereas those with weak or negative evaluations of both candidates would accurately express their attitudes by abstaining or casting a protest vote.

[17]We make the rating items mandatory because they capture respondents' evaluative preferences rather than their voting decisions. This ensures that even those who abstain or cast a protest vote still provide information about how they view each candidate. Practically, these ratings allow us to assess how the strength of candidate evaluations shapes respondents' likelihood of abstaining or opting out of choosing any candidate.

task includes one choice-based question and two rating-based questions assessing each candidate.

To ensure comparability, we replicate the hypothetical profile structures used in Hainmueller et al. (2014) for presidential elections and Peterson (2017) for congressional elections. These designs are ideal benchmarks for our study: Hainmueller et al. (2014) provides the canonical template for electoral conjoint experiments, and Peterson (2017) extends a similar structure to U.S. House races. Figure 4 Panel (a) presents an example of the candidate profiles shown in each conjoint task, and Panel (b) displays the corresponding choice questions administered to the control and treatment groups.

(a) Conjoint Task

Please carefully review the two candidates for US President detailed below. Which of the following two people do you think is more desirable as President of the United States?

	Candidate 1	**Candidate 2**
Religion	Evangelical Protestant	Mainline Protestant
Profession	High School Teacher	Farmer
Age	75	68
Annual Income	$54,000	$210,000
Race/Ethnicity	Caucasian	Black
Gender	Male	Male
Military Service	Served in U.S. military	No military service
College Education	BA from small college	BA from Baptist college

(b) Questions

Control Evaluation Questions

Which of these candidates would you vote to be **President of the United States**?

❑ Candidate 1
❑ Candidate 2

On a scale from 1 to 7, where 1 indicates that you would never support this candidate, and 7 indicates that you would always support this candidate, where would you place **Candidate 1**?

Never Support 1	2	3	4	5	6	**Definitely Support** 7
○	○	○	○	○	○	○

On a scale from 1 to 7, where 1 indicates that you would never support this candidate, and 7 indicates that you would always support this candidate, where would you place **Candidate 2**?

Never Support 1	2	3	4	5	6	**Definitely Support** 7
○	○	○	○	○	○	○

Treatment Evaluation Questions

Which of these candidates would you vote to be **President of the United States**?

❑ Candidate 1
❑ Candidate 2
❑ If I only have these two candidates, I will cast a blank/null vote.
❑ If I only have these two candidates, I will abstain.

On a scale from 1 to 7, where 1 indicates that you would never support this candidate, and 7 indicates that you would always support this candidate, where would you place **Candidate 1**?

Never Support 1	2	3	4	5	6	**Definitely Support** 7
○	○	○	○	○	○	○

On a scale from 1 to 7, where 1 indicates that you would never support this candidate, and 7 indicates that you would always support this candidate, where would you place **Candidate 2**?

Never Support 1	2	3	4	5	6	**Definitely Support** 7
○	○	○	○	○	○	○

Figure 4: Experimental Design: Presidential candidate conjoint study replicated based on Hainmueller et al. (2014). Panel (a) illustrates an example of profiles being compared in a task, while Panel (b) shows the questions that respondents in the control and treatment groups receive, respectively. Profiles and questions are presented on the same screen in each task for respondents.

We estimate a linear regression model that interacts candidate attributes with a treatment indicator to assess how AMCEs differ across forced- and unforced-choice designs. The outcome is coded as 1 for the selected profile, 0 for unselected or protest votes, and missing for abstentions. The model specification is detailed in Appendix B.

Designed-Induced Biases in Conjoint Estimates

In Conjoint Study 1, we replicate the candidate profiles from Hainmueller et al. (2014) in the context of a hypothetical U.S. presidential election. Respondents in the treatment group are given two additional options—casting a blank/null vote or abstaining—and may skip choice questions, whereas the control group must choose between Candidate 1 and Candidate 2. Among the 1,351 treated respondents, 391 cast at least one blank/null vote, 424 abstain at least once, and 671 use at least one opt-out option. Including skipped questions increases this to 689—over half the sample.

Most AMCEs remain stable across designs, but several differ meaningfully. For example, the AMCE for "Age 45 vs. Age 36" is 0.03 (SE = 0.01) under forced choice but shifts to –0.01 (SE = 0.01) in the unforced design. Although the absolute change is only –0.04, it fully reverses the effect's direction and reduces its magnitude by more than 100%. In the context of conjoint designs, where many estimated effects fall within a narrow numerical range, such shifts are nontrivial and can lead to different substantive conclusions.

Similar shifts are observed for other age levels (e.g., Age 52 and Age 68) and income levels (e.g., $54K, $65K, $210K vs. $32K). Figure 5 illustrates these differences, showing how the design influences the interpretation of attributes such as age and income; higher levels of income show no differences compared to the baseline in the treatment design. Interestingly, attributes that are insignificant in the control design become significant (e.g., Catholic vs. No Religion), while others lose significance (e.g., Income $5.1M vs. $32K). Taken together, these findings suggest that forced-choice designs may obscure meaningful patterns in voter behavior by overlooking actions such as abstention or protest voting, potentially distorting the interpretation of attribute effects.

Figure 5: AMCEs of Presidential candidate attributes, by design: The left and middle panels show the results for the Presidential candidates under the control and treatment designs. The rightmost panel shows the differences in AMCEs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level. The gray dots with gray bars indicate that there is no significance at either conventional level.

Additional Tests

To assess the generalizability of our findings, we replicate the analysis in a second conjoint experiment using hypothetical House candidate profiles from Peterson (2017), administered to the same sample. Results are presented in Figure A1. As in the first study, over half of respondents in the unforced-choice condition abstained or cast a protest vote at least once. While most AMCE estimates remain stable across designs, a few attribute effects shift slightly—generally smaller in magnitude and often significant at the 90% confidence level. Although some may view such changes as weak evidence, we interpret them differently. The selective nature of these differences indicates that expanding the outcome space to include abstention and protest voting does not inflate spurious design contrasts. Instead, the results suggest that design-induced bias is context-dependent: forced-choice designs do not always distort AMCE estimates, but when they do, the direction and magnitude of bias remain unpredictable ex ante.

As coding abstention as a deliberate missing value can drop observations and disrupt randomized balance, we assess robustness using both a two-step Heckman selection model and Inverse Probability Weighting (IPW), with results reported in Appendices C.2 and C.3. Across both approaches, the substantive conclusions remain unchanged.

In contrast, design-induced differences appear even more strongly in marginal means (MMs) as shown in Figures 6 and A2. MMs recover the absolute probability that a profile receives support (Leeper et al., 2020). This quantity can map onto the intended estimand: expected vote share when choices represent voting behavior, or simple appeal when choices reflect underlying candidate preferences. Forced-choice designs inflate support by requiring respondents to register support for one of the two profiles, thereby suppressing expressions of indifference or rejection. In the unforced-choice design, however, respondents who prefer neither profile can select abstention or a blank vote, reducing absolute support rates and revealing genuine non-preference.[18] As a result,

[18]In forced-choice designs, the grand mean is mechanically fixed at 0.5 because every task produces exactly one supported profile; distances from 0.5 are therefore interpreted as deviations from a "neutral" baseline of aggregate support. This convention, however, masks the possibility of respondent-level neutrality or rejection, since the design forces a choice even when neither profile is genuinely preferred. In unforced-choice designs, respondents may select neither option, so the grand mean is no longer anchored at 0.5. Marginal means are shaped both by underlying preferences and by the prevalence of meaningful non-choice behavior (e.g., blank votes), which enter as zeros for both profiles. As a result, marginal means in unforced-choice designs capture absolute support for each attribute level, without the artificial inflation created by forced choice. Accordingly, comparing favorability across attribute levels requires comparing the marginal means themselves—not their distance from 0.5.

MMs diverge more sharply across designs, which highlights that MMs are especially sensitive to how well the response options reflect the underlying behavioral choice being measured.

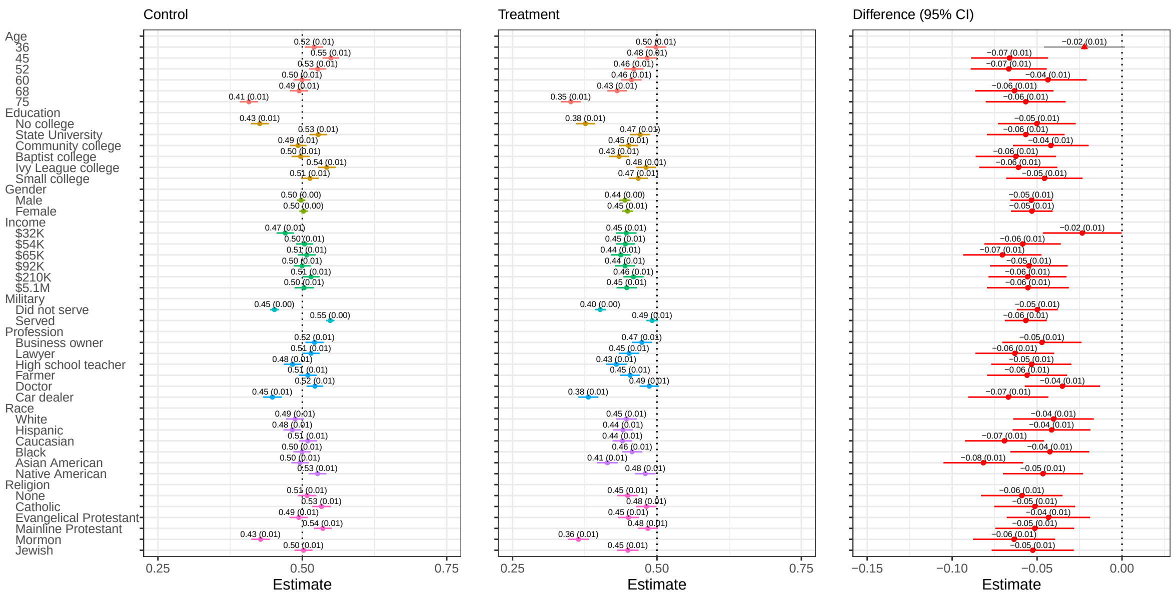

Figure 6: MMs of Presidential candidate attributes, by design: The left and middle panels show the results for the Presidential candidates under the control and treatment designs. The rightmost panel shows the differences in MMs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level.

We also address concerns about multiple comparisons. First, among the thirty-two estimated differences between the forced- and unforced-choice designs in Study 1, six are significant at the 5% level. While multiple comparison is a reasonable concern, if all thirty-two effects were truly null, the probability of observing six significant results by chance would be only about 0.37%.[19] While not impossible, such an chance is extremely small under the global null, suggesting that at least some of these differences reflect genuine design-induced effects. Second, as Liu and Shiraito (2023, 392) emphasize, "none of the (multiple testing correction) methods provides the perfect solution," because each entails a trade-off between Type I and Type II errors. A global correction across all conjoint coefficients would treat every attribute-level contrast as belonging to a single hypothesis family, imposing overly stringent thresholds. Importantly, all observed design-induced

[19]Under the global null, X $\sim$ Binomial(32, 0.05). $P(X=6)=\binom{32}{6}(0.05)^6(0.95)^{26}\approx 0.0037$.

differences occur within only two attributes (Age and Income), rather than being scattered across the eight attributes. A global approach risks obscuring real design-induced differences. We therefore define each attribute as its own hypothesis family and apply Benjamini–Hochberg adjustments within attributes. This targets the relevant source of multiplicity—coefficients that share a common baseline—while avoiding unnecessary conservatism. As shown in Figure A7, most adjusted p-values rise substantially, but several differences for age (45, 52, 68) and income ($65K) remain marginally significant, indicating that these effects are relatively robust to correction.

Lastly, a placebo test using the rating-based outcomes further addresses concerns about false positives. Unlike choice outcomes—which the forced-choice design directly constrains—ratings do not require respondents to select a candidate and therefore should be minimally affected by the treatment. Consistent with this expectation, and as shown in Figures A8 and A9, we find no significant differences between the forced- and unforced-choice groups on the 1–7 rating measures, with the exception of only one attribute-level in Conjoint Study 1.

Determinants of Switching to Opt-Out Options

To examine what drives respondents to opt out of choosing one of the two candidates in the unforced-choice design, we estimate linear regression models at the profile level for each replicated conjoint study. We recode the outcome as a binary indicator of opting out. In the first specification, any active choice—selecting Candidate 1, Candidate 2, or casting a blank vote—is coded as 0, and abstention is coded as 1. In the second specification, choosing either candidate is coded as 0, whereas selecting a blank or null (protest) vote or abstaining is coded as 1.[20] The models include all candidate attributes used in the conjoint design, along with pre-conjoint respondent-level covariates: past voting behavior, prospective turnout intentions, gender, age, education, race, partisanship, and ideological placement.[21] We also add task-level predictors, including the pair-

[20]Results are similar when comparing candidate choices with abstention only or with blank/null votes only, as shown in Figures A10 and A11.

[21]Exact question wordings are provided in Appendix D.

wise rating difference between the two profiles (0–6) and a contest-appeal indicator equal to 1 if both profiles are rated at or below 4. These predictors allow us to assess whether opting out is systematically linked to candidate-, respondent-, or task-level factors.

For Study 1, the results (Figures 7 and 8) show that shifts in respondent characteristics and task-level conditions exert much larger effects on opting out than do changes in any individual candidate attribute. Figure 7 shows that, holding all else constant, individuals who report voting "most of the time," "several times," or "seldom," compared to "always," exhibit progressively higher possibility of abstaining in the conjoint tasks. Respondents who report that they "will probably not vote" in the 2024 presidential election and who are older are also significantly more likely to abstain. At the task level, relative to tied pairs, each one-point increase in the rating difference between profiles in the same task decreases the likelihood of abstention,[22] whereas both profiles that are rated at or below 4 are strongly associated with greater abstention probabilities in the task.

However, almost all candidate attributes have little association with abstention. This suggests that changes in estimated effects between the two designs are not driven by imbalanced profile randomization when it is coded as missing values in the analysis of the design-induced differences in AMCE estimates but instead arise from the inclusion of expressive non-choice behaviors that forced-choice designs systematically suppress.

Figure 8 reports the determinants of opting out—either abstaining or casting a blank vote. After expanding the outcome to include blank votes, the effects of past voting frequency shift markedly: respondents who report voting in "most," "several," or "few" elections still show the monotonic rise in opt-out behavior observed in Figure 7, but these coefficients lose statistical significance. This pattern indicates that past abstention is not a strong driver of protest voting and that respondents clearly distinguish between abstaining and casting a null or blank ballot. Respondents who

[22]Because tasks in which both profiles are rated at or below 4 constrain the rating difference to less than 3, this variable can distort the estimated effect of larger rating differences (greater than 4). When we exclude the indicator for low-rated profile pairs, the relationship between rating difference and the likelihood of abstaining becomes progressively more negative, as expected.

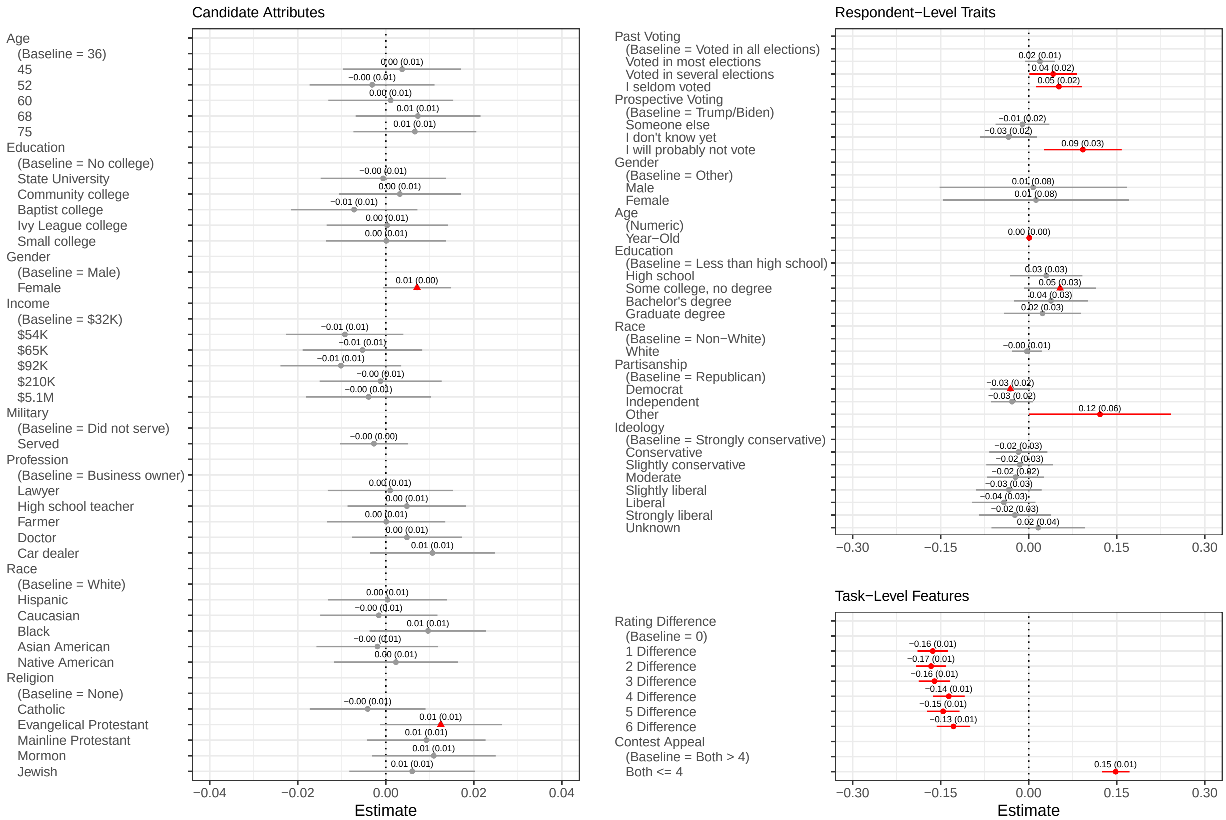

Figure 7: Determinants of Abstention in Conjoint Study 1. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of abstention: active choices (Candidate 1, Candidate 2, or blank) = 0; *abstention* = 1.Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

state that they "don't know yet" or "will probably not vote" in the 2024 presidential election are significantly more likely to opt out. Task-level predictors are even more highly associated with opting out: larger rating differences between the two profiles substantially reduce the likelihood of opting out, while tasks in which both profiles receive ratings of 4 or lower markedly increase it.

Most candidate-level attributes remain statistically insignificant, as in Figure 7. However, once blank votes are included in the outcome, a few attribute effects reach significance. This suggests that, while holding constant all other attributes, respondent characteristics, and task-level evaluations, certain candidate features can lead respondents to cast a blank or null vote rather than selecting either candidate.

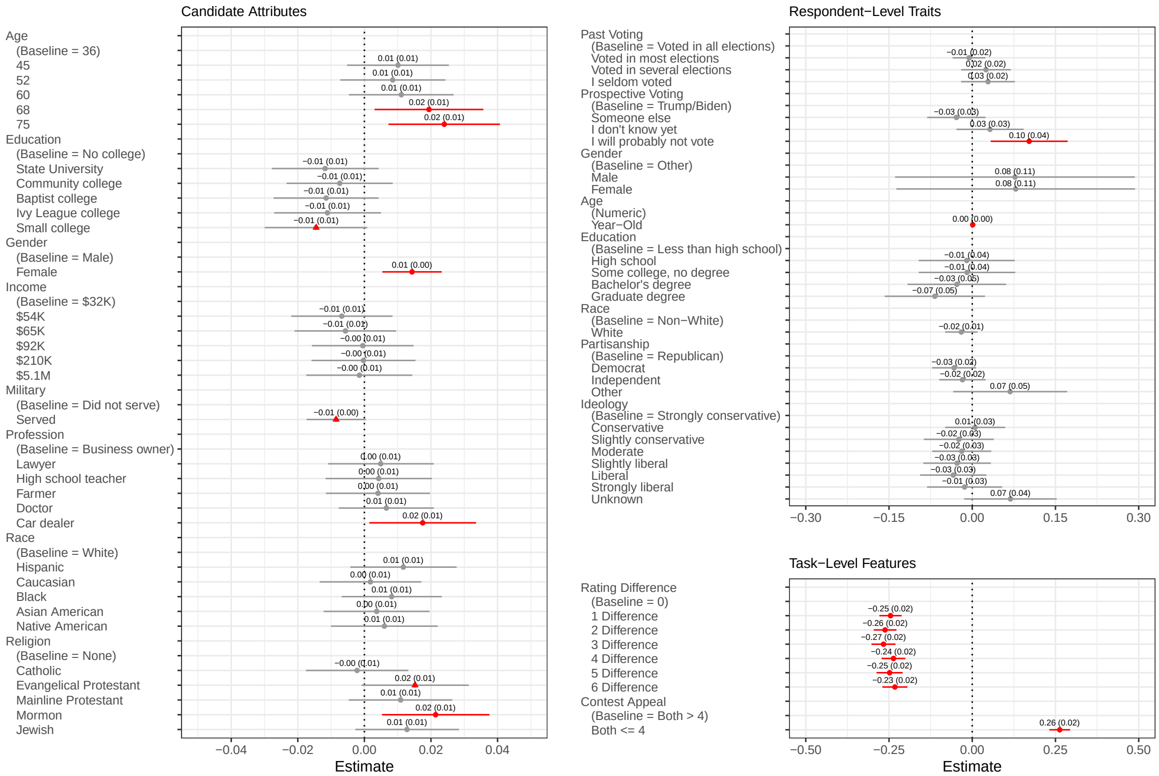

Figure 8: Determinants of Opting Out in Conjoint Study 1. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of opting out: choosing a candidate (Candidate 1 or Candidate 2) = 0; *selecting a blank vote or abstaining* = 1. Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

Study 2 shows a highly similar pattern, as illustrated in Figures A12 and A13. However, unlike Study 1, one candidate attribute shows a systematic association with opt-out behavior—abortion policy. Holding all else constant, respondents are more likely to abstain or submit a blank ballot when a candidate endorses the position that abortion should never be permitted. This indicates that certain salient specific policy stances can generate disaffection-driven opt-outs, even though respondent dispositions and contest-level evaluations remain the primary drivers.

Altogether, these results show that respondent- and contest-level factors systematically predict switching in outcome measures, challenging two core assumptions of forced-choice conjoint designs: that respondents must always select a preferred profile in each task and that they hold strict, complete preference orderings implying no indifference. A notable share of respondents—much as they did in real-world elections—opts for abstention or protest voting when given the opportunity. Moreover, the systematic patterns linking these behaviors to respondent- and task-level characteristics suggest that such non-selections are not random noise or inattentive behavior, but rather structured expressions of political preference. These results underscore that forced-choice designs, by their very nature, exclude a segment of genuine voter behavior—those who would otherwise abstain or cast a protest vote—thereby constraining the observed distribution of preferences and potentially biasing inferences about voter decision-making.[23]

Practice Guideline

Like other experimental designs, conjoint experiments rest on behavioral as well as statistical assumptions, yet the former are often left implicit. Forced-choice conjoint designs assume that respondents possess strict, complete preference orderings over all candidate profiles and that each respondent must make a selection. When estimates are interpreted as expected vote shares—or even as general preference patterns—these assumptions impose behavioral constraints rarely satisfied outside compulsory voting systems. Indifference or rejection toward both candidates is itself a meaningful preference. When respondents hold weak, indifferent, or negative views toward all presented options, binary forced-choice designs can generate measurement error and distort inferences about underlying preferences.

Researchers using forced-choice designs should therefore be explicit about these assumptions and cautious about the scope of their inferences. The format may be defensible in settings where turnout is legally mandated or normatively enforced, but in most democracies—where abstention, blank ballots, and "neither" responses are common political acts—extrapolating forced-choice

[23]We address two potential concerns—turnout self-overreporting and post-treatment bias—in Appendix C.7. Face-saving question wording yields a credible distribution of self-reported past turnout, with many respondents comfortably indicating low participation. Evidence from the control group shows that rating-based evaluations are not tightly anchored to the preceding choice, suggesting that any post-treatment bias in task-level predictors is limited.

findings to real electorates risks overstating participation and the clarity of preferences. For studies concerned primarily with preference intensity rather than expected vote shares, using rating-based outcomes provides a more flexible and behaviorally realistic measurement strategy than relying solely on forced binary choices.

When choice-based outcomes are essential and abstention or "none of the above" represents a meaningful dimension of voter choice, researchers should consider unforced-choice designs that explicitly include these options. Skipped questions do not distinguish deliberate abstention from inattentiveness, conflating disengagement with survey fatigue. Clear options such as "I would abstain" or "None of the above" provide respondents with an interpretable means of expressing disinterest or protest and preserve measurement validity.

The contrast between forced-choice and unforced-choice designs underscores a key trade-off in conjoint experiments. Forced-choice designs, at best, produce estimates comparable to those from unforced-choice designs, but they inevitably rest on two demanding behavioral assumptions that are often violated in practice. At worst, violating these assumptions can distort interpretation by conflating indifference or disapproval with genuine preference, masking meaningful variation in voter behavior. Unforced-choice designs relax these constraints by allowing respondents to opt out, capturing a full spectrum of political expression and yielding measures that more closely approximate real-world decision-making—even when the resulting estimates appear statistically similar. Researchers should therefore assess whether the behavioral assumptions embedded in forced-choice tasks align with their theoretical aims and empirical settings. When indifference or disengagement is substantively meaningful, incorporating explicit opt-out options is advisable.

Finally, regardless of the outcome format, we recommend that electoral conjoint studies include pre-conjoint questions capturing past turnout, turnout propensity, and key demographic and ideological traits. These measures help benchmark the sample against the broader electorate and illuminate heterogeneity in engagement. We advise against excluding respondents who report infrequent voting. Removing them introduces an additional—and often untenable—assumption that past abstention implies permanent nonparticipation. Instead, researchers should retain these respondents and allow explicit opt-out options to capture meaningful variation in political engagement.

Conclusion

Elections are a cornerstone of democracy, serving to select leaders, regulate competition, and resolve conflicts. Understanding voter decision-making is crucial but challenging, as traditional methods often overlook the multidimensional nature of choices and fail to replicate real-world behavior. Conjoint experiments address these limitations by enabling researchers to analyze the effects of multiple candidate attributes simultaneously (Hainmueller et al., 2014; Bansak et al., 2023).

This study revisits a foundational feature of most conjoint experiments in political science: the common usage of forced-choice designs. While this approach has become standard in studies of electoral preferences, it embeds two strong behavioral assumptions—that every respondent must choose one profile (Hainmueller et al., 2014) and that all respondents hold strict preferences over candidate profiles (Abramson et al., 2022; Bansak et al., 2023). These assumptions are rarely scrutinized but have important implications for interpretation and validity. When respondents harbor indifference, weak preferences, or disapproval toward all presented options, the forced-choice design imposes an artificial structure by constraining them to make a choice they may not endorse.

Our findings challenge the behavioral assumptions embedded in forced-choice conjoint designs. In our preregistered experiment, nearly half of respondents abstain or cast a protest vote at least once when given the opportunity—behavior that mirrors real-world turnout patterns and models of political disengagement. Although most AMCEs remain stable across designs, a subset shifts meaningfully in direction, magnitude, or significance. These changes are not random; they systematically correspond to respondent characteristics (such as past turnout and turnout intention) and task-level conditions (such as contest appeal and rating differentials). Forced-choice designs therefore risk compressing meaningful variation in political engagement, overstating participation and the strength of voters' preferences.

Substantively, the results speak directly to theories of political behavior and representation. First, they show that nonparticipation is itself a structured form of preference expression, not merely an absence of choice. Abstention and protest voting emerge as patterned political acts shaped by identifiable traits of voters and features of the choice environment. Second, our findings reveal that many citizens approach elections without strong or well-ordered preferences—information that standard forced-choice conjoint designs systematically obscure. This has direct implications for theories of voting behavior, expressive voting, and models of political accountability that assume consistent preference ordering. Finally, by demonstrating that opt-out behavior responds to contest appeal, our results contribute to broader debates on political disengagement, alienation, and the conditions under which citizens withdraw from electoral choice.

In settings where abstention, protest voting, or nonparticipation are common and politically meaningful, researchers should consider unforced-choice designs that explicitly incorporate opt-out options. When forced-choice designs are used, scholars should acknowledge the behavioral assumptions they impose and be cautious in generalizing their findings. Together, our methodological and theoretical contributions highlight that understanding what citizens choose not to do is just as informative and politically consequential as understanding the choices they make.

References

Abramson, S. F., K. Kocak, and A. Magazinnik (2022). What do we learn about voter preferences from conjoint experiments? *American Journal of Political Science 66*(4), 1008–1020.

Adams, J. and S. Merrill (2003). Voter turnout and candidate strategies in american elections. *The Journal of Politics 65*(1), 161–189.

Alvarez, R. M., D. R. Kiewiet, and L. Núñez (2018). A taxonomy of protest voting. *Annual Review of Political Science 21*(1), 135–154.

Bansak, K., J. Hainmueller, D. J. Hopkins, and T. Yamamoto (2018). The number of choice tasks and survey satisficing in conjoint experiments. *Political Analysis 26*(1), 112–119.

Bansak, K., J. Hainmueller, D. J. Hopkins, and T. Yamamoto (2023). Using conjoint experiments to analyze election outcomes: The essential role of the average marginal component effect. *Political Analysis 31*(4), 500–518.

Battaglini, M., R. B. Morton, and T. R. Palfrey (2008). Information aggregation and strategic abstention in large laboratory elections. *The American Economic Review 98*(2), 194–200.

Bernhagen, P. and M. Marsh (2007). The partisan effects of low turnout: Analyzing vote abstention as a missing data problem. *Electoral Studies 26*(3), 548–560.

Blais, A. and J.-F. Daoust (2020). *The Motivation to Vote: Explaining Electoral Participation*. University of British Columbia Press. ProQuest Ebook Central.

Clayton, K., Y. Horiuchi, A. R. Kaufman, G. King, and M. Komisarchik (2025). Correcting measurement error bias in conjoint survey experiments. *American Journal of Political Science*. Forthcoming.

Cohen, M. J. (2018). Protesting via the null ballot: An assessment of the decision to cast an invalid vote in latin america. *Political Behavior 40*(2), pp. 395–414.

de la Cuesta, B., N. Egami, and K. Imai (2022). Improving the external validity of conjoint analysis: The essential role of profile distribution. *Political Analysis 30*(1), 19–45.

Degtiar, I. and S. Rose (2023). A review of generalizability and transportability. *Annual Review of Statistics and Its Application 10*(Volume 10, 2023), 501–524.

Eggers, A. C., N. Vivyan, and M. Wagner (2018). Corruption, accountability, and gender: Do female politicians face higher standards in public life? *The Journal of Politics 80*(1), 321–326.

Fu, J. and X. Li (2025). Generalization issues in conjoint experiment: Attention and salience.

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

Hainmueller, J., D. Hangartner, and T. Yamamoto (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences 112*(8), 2395–2400.

Hainmueller, J., D. J. Hopkins, and T. Yamamoto (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis 22*(1), 1–30.

Horiuchi, Y., Z. Markovich, and T. Yamamoto (2021). Does conjoint analysis mitigate social desirability bias? *Political Analysis 30*(4), 535–549.

Horiuchi, Y., D. M. Smith, and T. Yamamoto (2020). Identifying voter preferences for politicians' personal attributes: a conjoint experiment in japan. *Political Science Research and Methods 8*(1), 75–91.

Kane, J. V. and M. Costa (2024). Being careful with conjoints: Accounting for inattentiveness in conjoint experiments. APSA Preprints, October 11, 2024. DOI: 10.33774/apsa-2024-xqt7d-v2.

Kawai, K., Y. Toyama, and Y. Watanabe (2021, November). Voter turnout and preference aggregation. *American Economic Journal: Microeconomics 13*(4), 548–86.

Kirkland, P. A. and A. Coppock (2018, September). Candidate Choice Without Party Labels:. *Political Behavior 40*(3), 571–591.

Koch, C. M., C. Meléndez, and C. Rovira (2023). Mainstream voters, non-voters and populist voters: What sets them apart? *Political Studies 71*(3), 893–913.

Leeper, T. J., S. B. Hobolt, and J. Tilley (2020). Measuring subgroup preferences in conjoint experiments. *Political Analysis 28*(2), 207–221.

Liu, G. and Y. Shiraito (2023). Multiple hypothesis testing in conjoint analysis. *Political Analysis 31*(3), 380–395.

Meyer, B. D. and N. Mittag (2017). Misclassification in binary choice models. *Journal of Econometrics 200*(2), 295–311.

Miller, D. R. and J. Ziegler (2024). Preferential abstention in conjoint experiments. *Research & Politics 11*(4), 1–8.

Peterson, E. (2017). The role of the information environment in partisan voting. *The Journal of Politics 79*(4), 1191–1204.

Plescia, C., S. Kritzinger, and S. P. Singh (2023). Who would vote nota? explaining a 'none of the above' choice in eight countries. *European Journal of Political Research 62*(1), 118–134.

Shu, D. and G. Y. Yi (2019). Causal inference with measurement error in outcomes: Bias analysis and estimation methods. *Statistical Methods in Medical Research 28*(7), 2049–2068.

Smets, K. and C. van Ham (2013). The embarrassment of riches? a meta-analysis of individual-level research on voter turnout. *Electoral Studies 32*(2), 344–359.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Zhirkov, K. (2022). Estimating and using individual marginal component effects from conjoint experiments. *Political Analysis 30*(2), 236–249.

Supplementary Information (Online)

The Limitations of Using Forced Choice in Electoral Conjoint Experiments

Contents

A List of Published Articles Using Conjoint Analysis in Electoral Studies

We focus on articles published or accepted in ten leading political science journals between 2014 and 2023: American Journal of Political Science (AJPS), American Political Science Review (APSR), British Journal of Political Science (BJPS), Comparative Political Studies (CPS), Electoral Studies (Elect. Stud.), Journal of Politics (JOP), Political Behavior (PolBeh), Public Opinion Quarterly (POQ), and Political Science Research and Methods (PSRM).

Paper	Journal	Context of Study	Forced-Choice Design
Thomas et al (2024)	AJPS	India	Y
Greene (2022)	AJPS	Mexico	Y
Spater (2022)	AJPS	India	Y
Costa (2021)	AJPS	US	Y
Frederiksen (2022)	APSR	Many	Y
Teele, Kalla and Rosenbluth (2018)	APSR	US	Y
Carnes and Lupu (2016)	APSR	US, UK, and Argentina	Y
Rains and Wibbels (2023)	BJPS	India	Y
Loewen and Rheault (2021)	BJPS	US and Canada	Y
Campbell et al (2019)	BJPS	UK	Y
Schuler (Forthcoming)	CPS	Vietnam	Y
Laterzo (2024)	CPS	Argentina and Brazil	Y
Ventura, Ley and Cantú (2023)	CPS	Mexico	Y
Hankla et al (2023)	CPS	India	Y
Singh (2022)	CPS	Argentina	Y
Chou et al (2021)	CPS	Germany	Y
Carter (2021)	CPS	Peru	Y
Clayton et al (2019)	CPS	Malawi	Y
Agerberg (2019)	CPS	Spain	N
Arceneaux and Wielen (2023)	Elect. Stud.	US	Y
Busby (2022)	Elect. Stud.	US	Y
Kim (2021)	Elect. Stud.	Kenya	Y
Kang et al (2021)	Elect. Stud.	Australia	Y
Foulon and Reyes-Housholder (2021)	Elect. Stud.	Uruguay, Argentina and Chile	Y
Shockley and Gengler (2020)	Elect. Stud.	Qatar	Y
Vivyan et al (2020)	Elect. Stud.	Austria, Germany and UK	Y
Badas and Stauffer (2019)	Elect. Stud.	US	Y
Arnesen, Duell and Johannesson (2019)	Elect. Stud.	Norway	Y
Gift and Lastra-Anadón (2018)	Elect. Stud.	US	Y
Matsuo and Lee (2018)	Elect. Stud.	UK	Y
Aguilar, Cunow and Desposato (2015)	Elect. Stud.	Brazil	Y

Continued on next page

Paper	**Journal**	**Context of Study**	**Forced-Choice Design**
Frederiksen (2024)	JOP	US, UK, Czech, Mexico, and South Korea	Y
Driscoll and Nelson (2023)	JOP	US	Y
Portmann (2022)	JOP	Switzerland	Y
Henderson et al (2022)	JOP	US	Y
Eshima and Smith (2022)	JOP	Japan	Y
Weaver (2021)	JOP	Peru	Y
Magni and Reynolds (2021)	JOP	US, UK, and New Zealand	Y
Bakker, Schumacher and Rooduijn (2021)	JOP	US	Y
Schneider (2020)	JOP	Germany	Y
Chauchard, Klašnja and Harish (2019)	JOP	India	Y
Campbell et al (2019)	JOP	UK	Y
Peterson and Simonovits (2018)	JOP	US	Y
Eggers, Vivyan and Wagner (2018)	JOP	UK	N
Peterson (2017)	JOP	US	Y
Robinson (2023)	PolBeh	US	Y
Wood (2023)	PolBeh	US	Y
Visconti (2022)	PolBeh	Chile	Y
Magni and Reynolds (2022)	PolBeh	US, UK, and New Zealand	Y
Manento and Testa (2022)	PolBeh	US	Y
Rehmert (2022)	PolBeh	Germany	Y
Funck and McCabe (2022)	PolBeh	US	Y
Neuner and Wratil (2022)	PolBeh	Germany	Y
Saha and Weeks (2022)	PolBeh	US and UK	Y
Martin and Blinder (2021)	PolBeh	UK	Y
Rosenzweig (2021)	PolBeh	Kenya	Y
Blackman and Jackson (2021)	PolBeh	Tusnia	Y
Mummolo, Peterson and Westwood (2021)	PolBeh	US	Y
Crowder-Meyer et al (2020)	PolBeh	US	Y
Leeper and Robison (2020)	PolBeh	US	Y
Ono and Burden (2019)	PolBeh	US	Y
Kirkland and Coppock (2018)	PolBeh	US	Y
Sances (2018)	PolBeh	US	Y
DeMora et al (2022)	POQ	US	Y
Lehrer, Stöckle and Juhl (2024)	PSRM	Germany	N
Dai and Kustov (2023)	PSRM	US	Y
Erlich and Beauvais (2023)	PSRM	Ukraine	N
Ono and Yamada (2020)	PSRM	Japan	Y
Mares and Visconti (2020)	PSRM	Romania	Y
Horiuchi, Smith and Yamamoto (2020)	PSRM	Japan	Y
Franchino and Zucchini (2015)	PSRM	Italy	Y
Carlson (2015)	WP	Uganda	Y

B Model Specification

We define a sample of respondents indexed by *i* (i=1,..., N), and each respondent evaluates a predefined number of profiles indexed by j (j=1,..., J). The outcome variable y_{ij}, which is part of the vector (y_{i1}, ..., y_{iJ}) represents either choices or ratings (on a scale from 0 to 7) given by respondent i to the profile j presented. Profiles are described in terms of attributes indexed by *k* (k=1,..., K). $\mathbf{X}_{ik,j}$ is the vector of profile attributes, each element of which is a *factor variable* describing a certain attribute value presented to respondent i in profile j. The variable D_i is a binary variable that indicates the type of conjoint design randomly assigned at the individual level. It takes the value of 1 if the treatment design is assigned, and 0 otherwise. Then, the regression takes the following form:

$$y_{ij} = \sum_{k=1}^{K} \alpha_k \mathbf{X}_{ik,j} + \beta D_i + \sum_{k=1}^{K} \lambda_k \mathbf{X}_{ik,j} \times D_i + \varepsilon_{ij} \tag{1}$$

where α_k, β, and λ_k are regression parameters to be estimated, and ε_{ij} is the respondent-profile-specific error term. We are particularly interested in the size and significance of the estimate for parameter λ_k. This indicates the size and direction of the AMCEs observed from the treatment design different to the control design. We code y_{ij} based on the design assignments. For respondent i who received the typical electoral conjoint design (control design), y_{ij} is coded as 1 if profile j is selected and 0 if it is not. In contrast, for respondent i who received our proposed electoral conjoint design (treatment design), y_{ij} is coded as follows: if respondent i selects either Candidate A or B, the chosen profile is coded as 1 and the unchosen profile as 0. If respondent i casts a protest vote, both profiles are coded as 0. If abstention is chosen or the question is skipped, both profiles are coded as missing values.

C Additional Tests and Discussion

C.1 Conjoint Study 2

In Conjoint Study 2, with the same respondent sample, we replicated the profiles of hypothetical congressional candidates from Peterson (2017). Although respondents complete both studies in a random order, they remain with the same treatment assignment throughout the entire survey. The treatment design is same with that in Study 1, allowing respondents to abstain or cast blank/null votes. Of the 1,351 respondents in the treatment group, 427 cast at least one protest vote, 479 abstained at least once, and 741 (over half) opted for either abstention or a protest vote. Following Peterson (2017), each conjoint task paired a Democrat and a Republican candidate, mirroring a typical general-election matchup.

As in Study 1, AMCE estimates differed between the treatment and control designs, although the differences in Study 2 were typically smaller and significant at the 90% confidence level. Figure A1 highlights these shifts, showing how the inclusion of real-world voting options affects AMCE estimates and their interpretation. For example, the forced-choice design suggested that candidates aged 46 had a higher likelihood of being selected, while candidates aged 74 were disadvantaged compared to those aged 28. The treatment design, however, indicated no significant age-related advantage. Similarly, while the forced-choice design suggested a significant advantage for married candidates (married or divorced) over single ones, the treatment design showed that marital status was not a significant predictor of voter preference. Conversely, the treatment design revealed a statistically significant preference for female candidates, which was not observed in the forced-choice design.

As a note, readers might observe larger differences between the treated and control conditions when comparing the presidential to the congressional conjoint experiments. A plausible explanation for this pattern could be that the first conjoint included multiple relevant attributes, while the second conjoint featured one very salient attribute for all types of voters (i.e., abortion). However, since researchers cannot predict in advance which attribute will primarily drive the results, we recommend always including alternative options, such as abstaining or casting null/blank ballots.

The marginal-mean results from Conjoint Study 2 display the same broad pattern observed in the presidential conjoint: unconditional support levels are substantially lower in the unforced-choice design across nearly all attribute levels, with differences often ranging between 3 and 8 percentage points. As shown in A2, forced-choice tasks mechanically inflate marginal means by requiring respondents to select one of the two candidates, whereas the unforced-choice design allows disaffected or indifferent respondents to register non-support. These patterns reinforce that marginal means, because they capture absolute support, are especially sensitive to the outcome space and will diverge more sharply when forced-choice assumptions are relaxed.

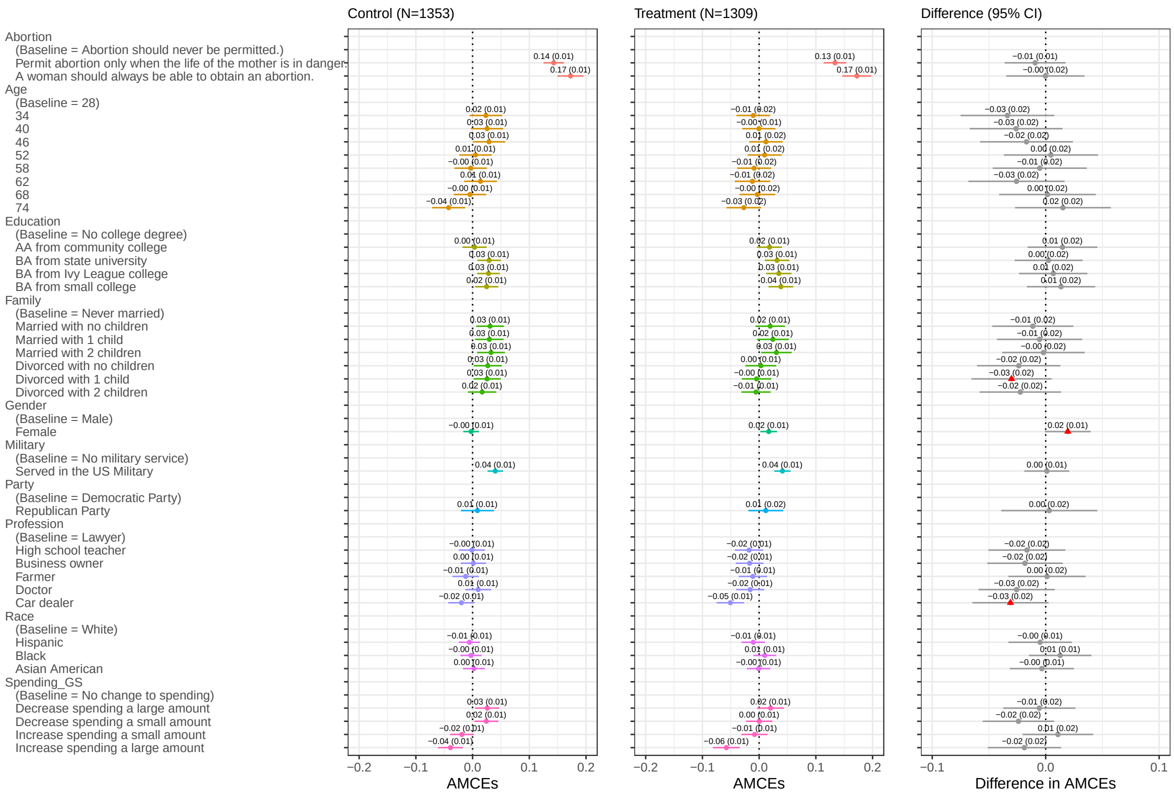

Figure A1: AMCEs of Congressional candidate attributes by design: The left and middle panels show the results for the Congressional candidates under the control and treatment designs. The rightmost panel shows the differences in AMCEs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level. The gray dots with gray bars indicate that there is no significance at either conventional level.

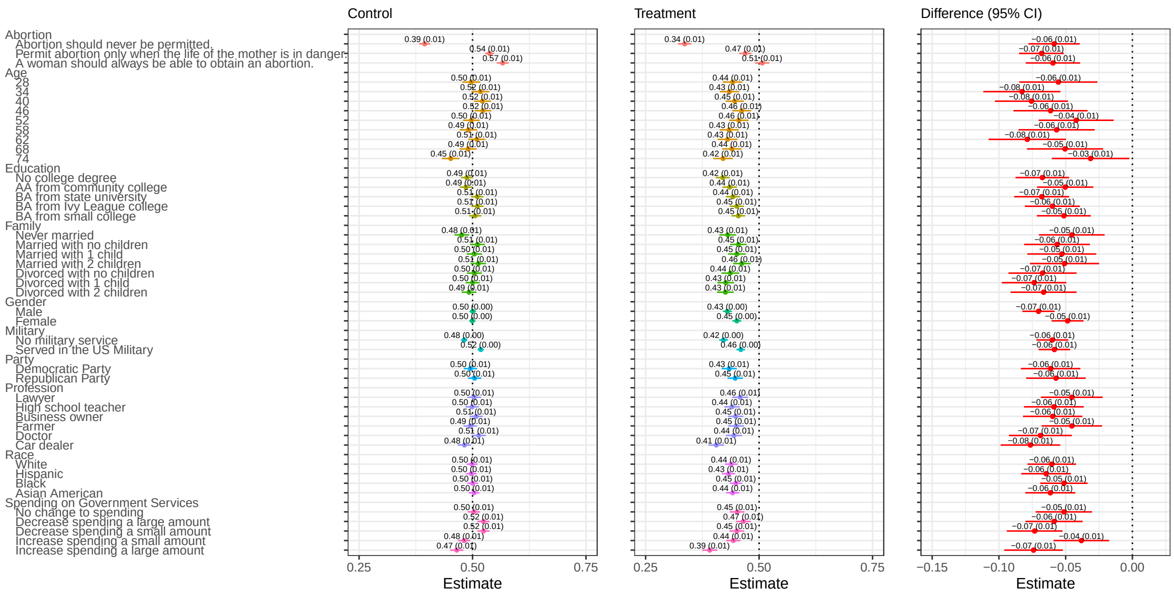

Figure A2: MMs of Congressional candidate attributes, by design: The left and middle panels show the results for the Congressional candidates under the control and treatment designs. The rightmost panel shows the differences in MMs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level.

C.2 Two-Step Heckman Selection

We first employ the Two-Step Heckman Selection Approach, which is designed to account for selection bias arising from non-random missingness. The Heckman approach consists of two stages. The first stage, known as the selection equation, models the likelihood that an observation is included in the sample, which in this case corresponds to whether the respondent chooses to turn out ($T_i = 1$). We use a probit model for this stage, predicting the probability of turnout based on both profile attributes and respondent-level pre-treatment covariates. The predictors in the selection equation include candidate or policy features presented in the conjoint task, as well as respondent characteristics such as past voting history, propensity to vote in the upcoming election, age, education, income, gender, and race. Mathematically, the selection equation is expressed as:

$$\Pr(T_i = 1) = \Phi(Z_i\beta) \tag{2}$$

where $T_i = 1$ indicates the respondent turns out, Z_i includes profile attributes and respondent covariates, β represents the coefficients, and Φ is the cumulative distribution function of the probit model.

From this stage, we compute the inverse Mills ratio (IMR) for each observation, which captures the likelihood of inclusion based on the selection process. The IMR is given by:

$$\text{IMR}_i = \frac{\phi(Z_i\beta)}{\Phi(Z_i\beta)} \tag{3}$$

where ϕ and Φ are the probability density function and cumulative density function of the normal distribution, respectively. This ratio is later used in the second stage to adjust for the selection bias. The second stage, known as the outcome equation, models the relationship between candidate attributes (treatment variables) and the outcome variable (Y_i) for the subset of respondents who chose to turn out. This stage is a linear regression that resembles that in a conjoint analysis but includes the inverse Mills ratio as an additional explanatory variable to control for selection bias. The outcome equation is expressed as:

$$Y_i = X_i\gamma + \lambda \cdot \text{IMR}_i + \epsilon_i, \tag{4}$$

where Y_i is the choice-based outcome variable (e.g., vote choice), X_i represents the treatment variables (candidate attributes), γ are the coefficients for the treatment effects, λ is the coefficient for the inverse Mills ratio, and ϵ_i is the clustered error term. The inclusion of the IMR ensures that the estimates of γ are unbiased, even in the presence of selection bias. The coefficients in the outcome equation represent the effects of candidate attributes on the outcome variable, conditional on respondents being included in the sample (i.e., having chosen to turn out). The selection equation accounts for the factors influencing the decision to abstain, thereby correcting for bias introduced by non-random missingness. By explicitly modeling the selection process, the Two-Step Heckman Selection Approach provides robust estimates of the causal effects of candidate attributes, even when abstention behavior leads to deviations from complete randomization. This approach ensures that our analysis remains valid and that the observed treatment effects are accurately estimated, as shown in Figures A3 and A4.

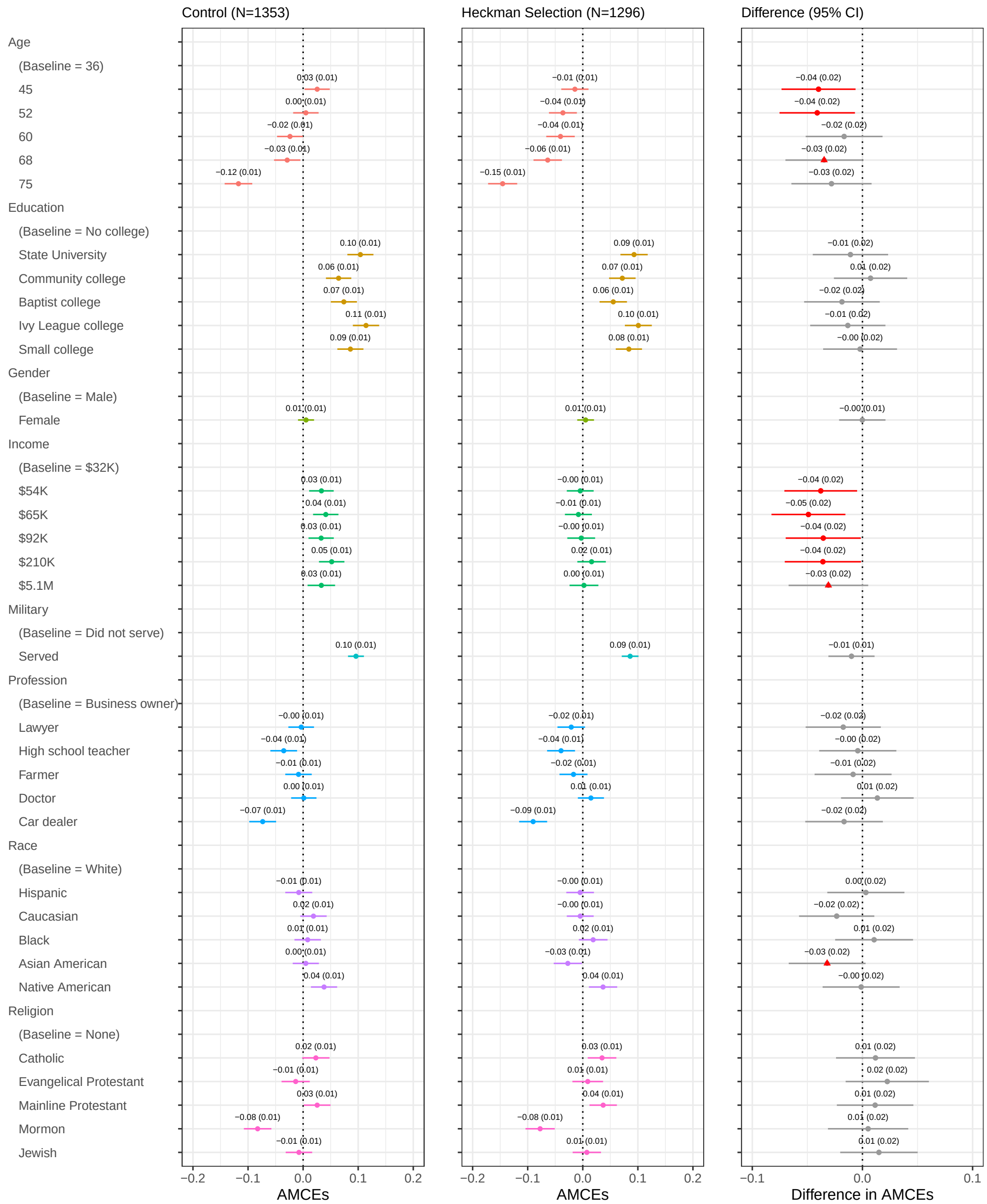

Figure A3: AMCEs of Presidential Candidate Attributes by Design Using Heckman-Selection Model: The left and middle panels present the AMCE results for Presidential candidates under the control and treatment designs. The treatment design estimates were obtained using the 2-step Heckman selection. The rightmost panel displays the differences in AMCEs for each attribute level, with horizontal bars representing 95% confidence intervals robust to clustering at the respondent level. Red dots with red bars denote significance at the 5% level, while red triangles with gray bars indicate significance only at the 10% level. Gray dots with gray bars represent attribute levels that are not significant at either conventional level.

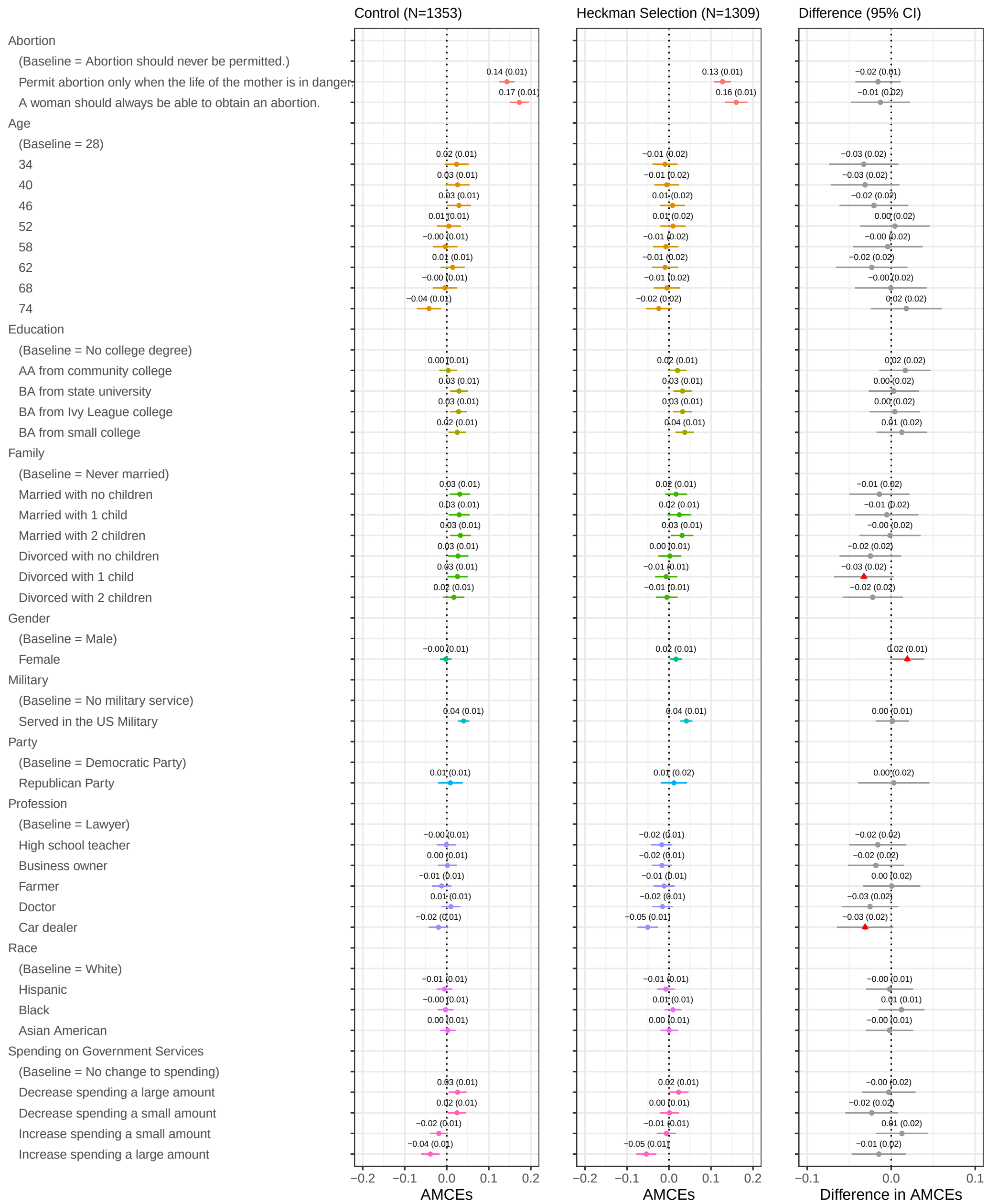

Figure A4: AMCEs of Congressional Candidate Attributes by Design Using Heckman-Selection Model: The left and middle panels present the AMCE results for Congressional candidates under the control and treatment designs. The treatment design estimates were obtained using the 2-step Heckman selection. The rightmost panel displays the differences in AMCEs for each attribute level, with horizontal bars representing 95% confidence intervals robust to clustering at the respondent level. Red dots with red bars indicate significance at the 5% level, while red triangles with gray bars denote significance only at the 10% level. Gray dots with gray bars represent attribute levels that are not significant at either conventional level.

C.3 Inverse Probability Weighting

We also apply Inverse Probability Weighting (IPW) as an alternative approach to mitigate potential selection bias. Unlike the Two-Step Heckman Selection model, which explicitly models the selection process, IPW adjusts for bias by reweighting the observed data to approximate the distribution of the target population.

The IPW approach begins by estimating the probability of each respondent turning out ($T_i = 1$), conditional on both candidate attributes and respondent-level pre-treatment covariates. A logistic regression model is used to estimate the turnout probability:

$$\Pr(T_i = 1 | D_j^{(k)}, X_i) = \Phi(Z_i \beta) \tag{5}$$

where Z_i includes the profile attributes ($D_j^{(k)}$) and individual covariates (X_i), and Φ represents the cumulative distribution function of the logistic model. Once the probabilities are estimated, weights are assigned to each observation as the inverse of their estimated probability of turnout:

$$w_i = \frac{1}{\Pr(T_i = 1 | D_j^{(k)}, X_i)} \tag{6}$$

These weights are inversely proportional to the likelihood of being observed in the sample. Observations with a lower probability of turnout (based on their characteristics and the presented profiles) are assigned higher weights, ensuring that the weighted sample better represents the overall target population.

In the outcome analysis, these weights are incorporated directly into the estimation of the Average Marginal Component Effects (AMCEs). The weighted AMCE is calculated as:

$$\text{AMCE}j^{(kk\prime)} = \frac{\sum_{i=1}^{N} w_i \cdot (Y_i | D_j^{(k)} = 1)}{\sum_{i=1}^{N} w_i} - \frac{\sum_{i=1}^{N} w_i \cdot (Y_i | D_j^{(k\prime)} = 1)}{\sum_{i=1}^{N} w_i} \tag{7}$$

This weighted estimator adjusts for the non-random selection process introduced by abstention, ensuring that the estimated treatment effects are robust to selection bias.

The key advantage of IPW lies in its flexibility. By reweighting the observed data, it addresses differences in the likelihood of turnout across subgroups without requiring explicit modeling of the selection process in the outcome equation. Additionally, using both pre-treatment covariates and candidate attributes in the weighting model ensures that the estimated probabilities capture the primary factors influencing abstention, reducing bias and aligning the weighted study population with the intended target population. Figures A5 and A6 illustrate the robustness of our results when using this approach.

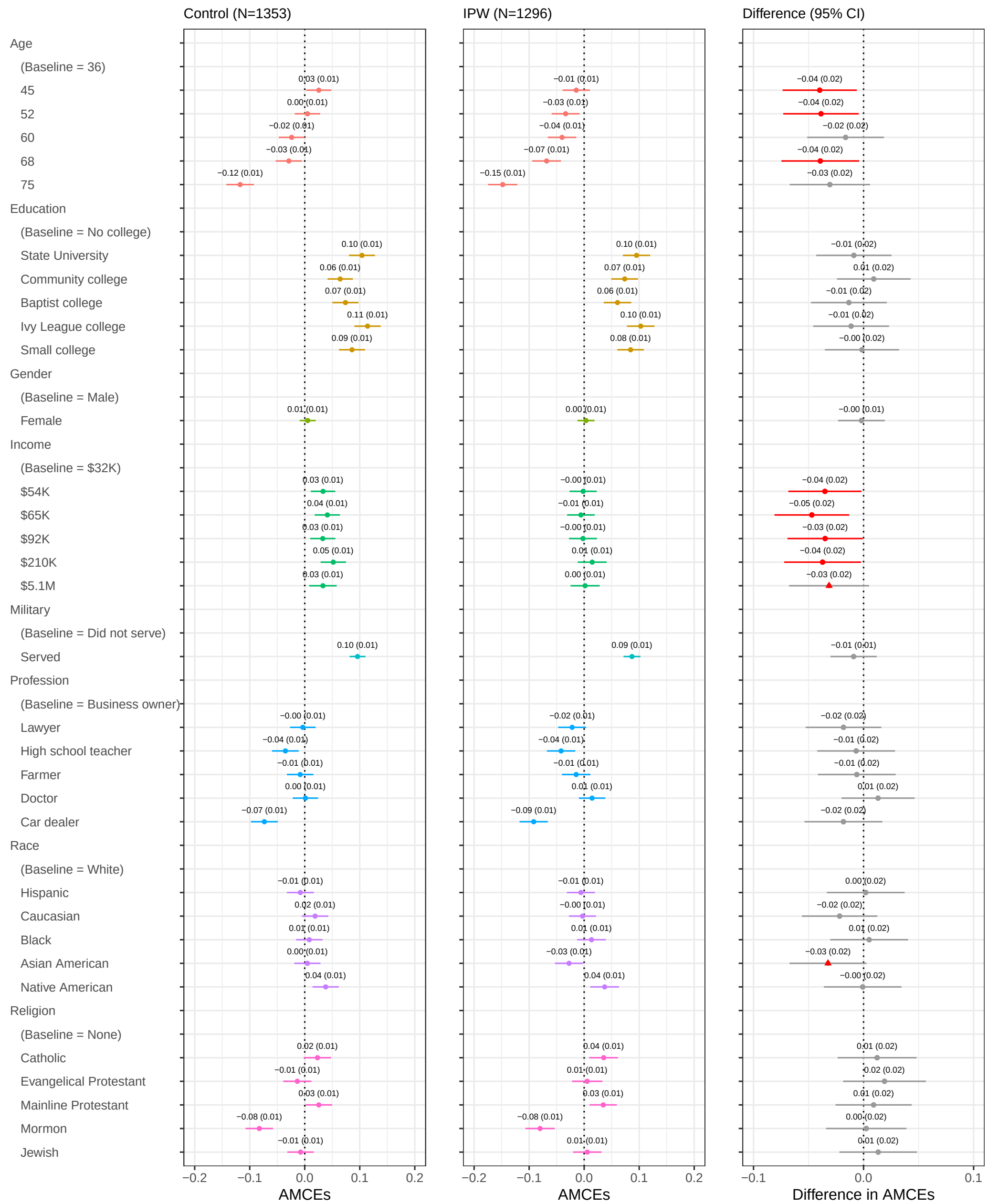

Figure A5: AMCEs of Presidential Candidate Attributes by Design Using IPW: The left and middle panels present the AMCE results for Presidential candidates under the control and treatment designs. The treatment design estimates were obtained using the Inverse Probability Weighting. The rightmost panel displays the differences in AMCEs for each attribute level, with horizontal bars representing 95% confidence intervals robust to clustering at the respondent level. Red dots with red bars denote significance at the 5% level, while red triangles with gray bars indicate significance only at the 10% level. Gray dots with gray bars represent attribute levels that are not significant at either conventional level.

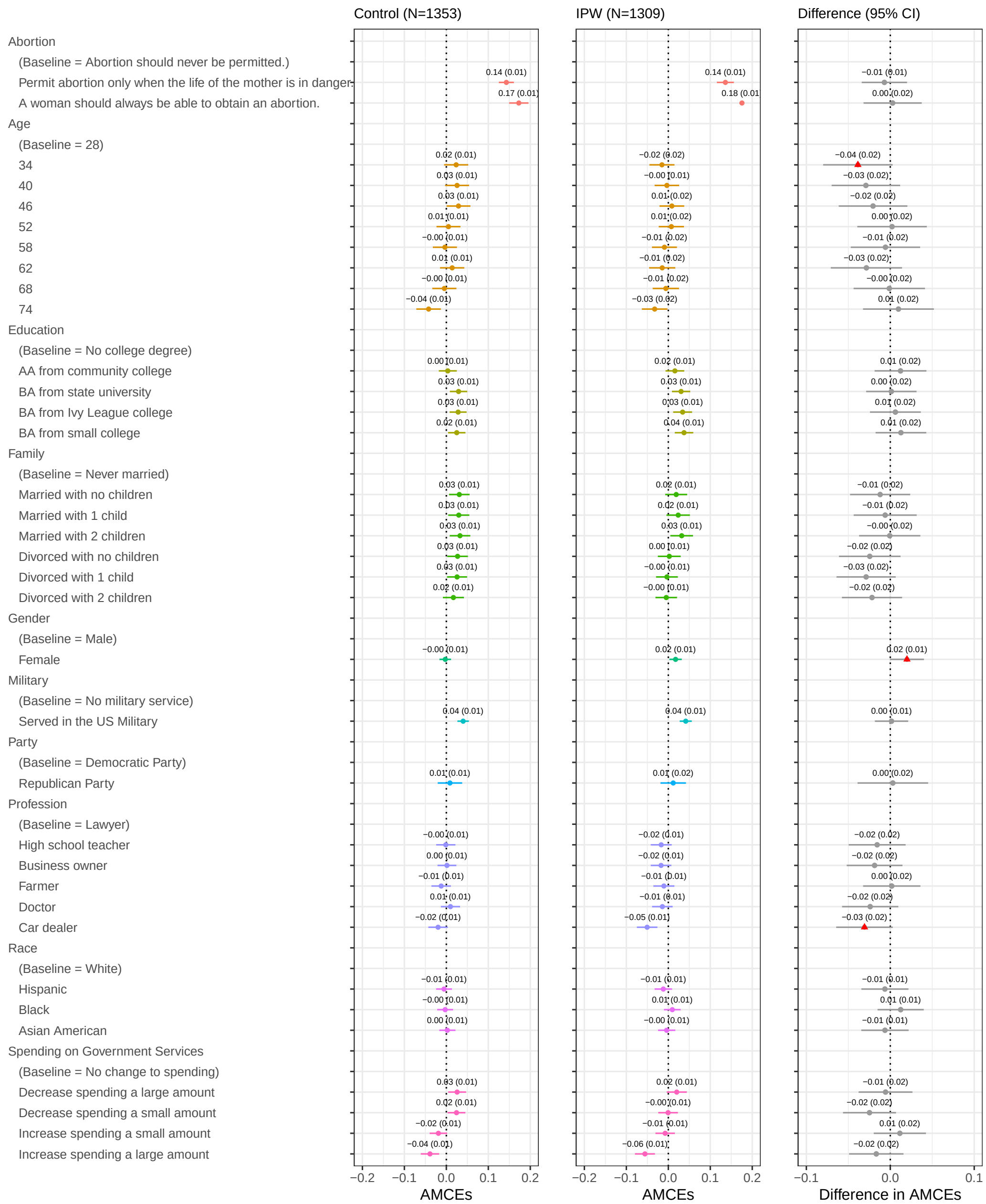

Figure A6: AMCEs of Congressional Candidate Attributes by Design Using IPW: The left and middle panels present the AMCE results for Congressional candidates under the control and treatment designs. The treatment design estimates were obtained using the Inverse Probability Weighting. The rightmost panel displays the differences in AMCEs for each attribute level, with horizontal bars representing 95% confidence intervals robust to clustering at the respondent level. Red dots with red bars indicate significance at the 5% level, while red triangles with gray bars denote significance only at the 10% level. Gray dots with gray bars represent attribute levels that are not significant at either conventional level.

C.4 Multiple-Testing Adjustments

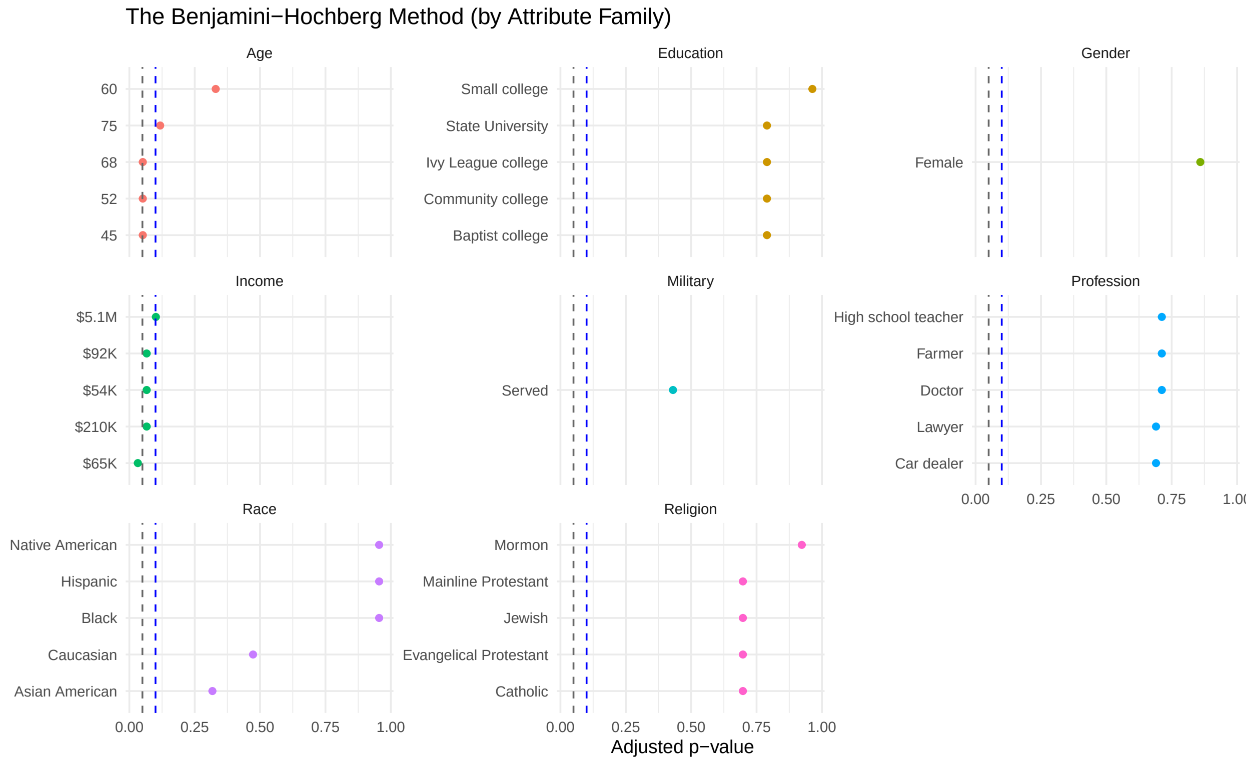

Figure A7: Benjamini–Hochberg Adjusted p-Values by Attribute in Conjoint Study 1. Adjusted p-values from the Benjamini–Hochberg (BH) false discovery rate correction, applied within each attribute family. The grey dotted line marks the conventional significance threshold (p=0.05), and the blue dotted line marks the marginal significance threshold (p=0.1).

C.5 Placebo Test: Design-Induced Differences in Rating-Based Estimates

To assess whether the design-induced differences in the choice outcomes could be artifacts of multiple testing rather than genuine behavioral responses, we conduct a placebo test using the 1–7 rating items. These ratings capture respondents' preference intensity of each candidate. Crucially, unlike the choice-based outcome—which the forced-choice design directly constrains by requiring respondents to select one of the two candidates—the rating questions do not force a choice and should therefore be largely unaffected by the treatment, even if some respondents might naturally align their ratings with their choices. If the significant differences observed in the choice outcomes were driven by random noise or inflated false positives, we would expect similar spurious differences in the rating outcomes.

We estimate rating-based AMCEs for the forced-choice and unforced-choice groups using the same attribute-level contrasts as in the main analysis. As shown in Figures A8 and A9, we find virtually no significant differences across conditions; only one attribute-level, Age 45, contrast reaches significance. Moreover, when we re-estimate the unforced-choice AMCEs using ratings as the outcome, the resulting effects closely mirror those obtained from the choice-based estimands.

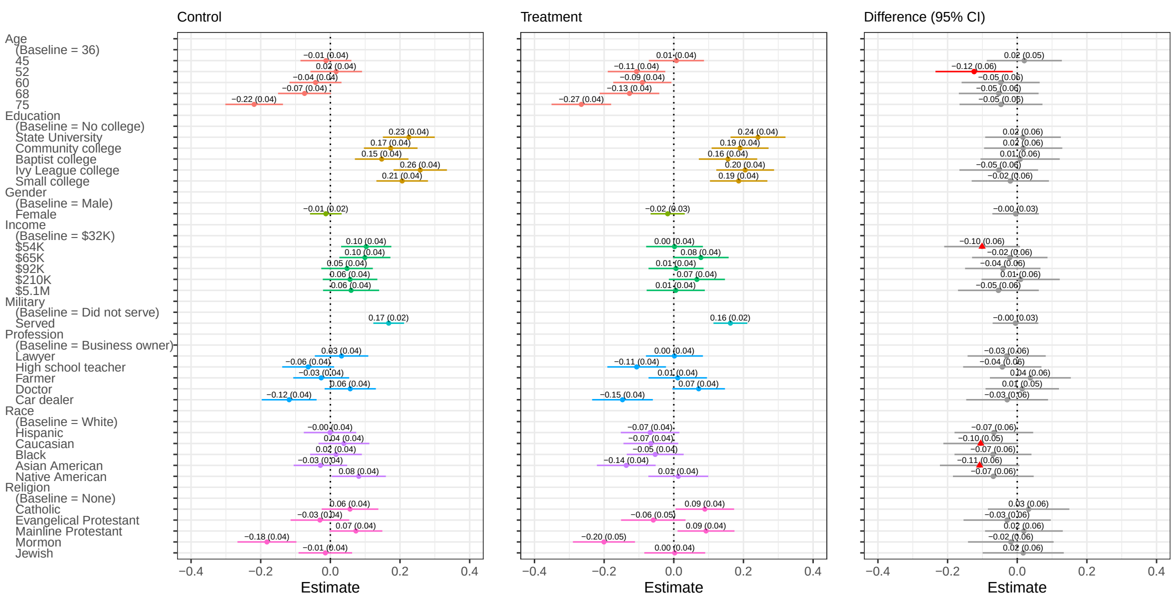

Figure A8: Rating-based AMCEs of Presidential candidate attributes by design: The left and middle panels show the results for the Presidential candidates under the control and treatment designs. The rightmost panel shows the differences in AMCEs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level. The gray dots with gray bars indicate that there is no significance at either conventional level.

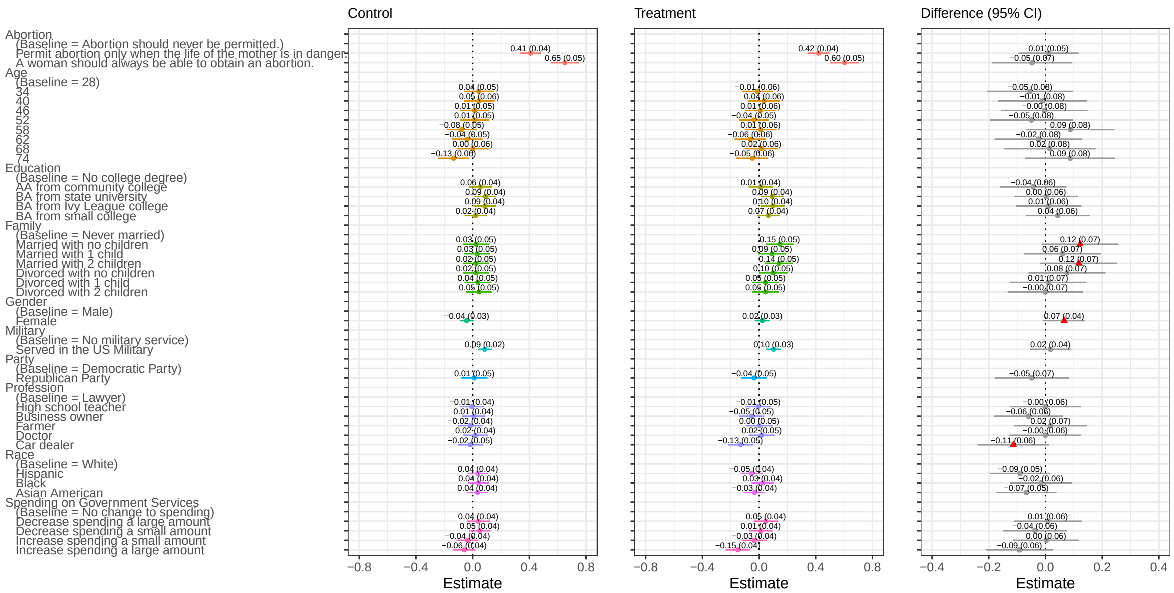

Figure A9: Rating-based AMCEs of Congressional candidate attributes by design: The left and middle panels show the results for the Congressional candidates under the control and treatment designs. The rightmost panel shows the differences in AMCEs for each attribute level, in which horizontal bars represent 95% confidence intervals robust to clustering at the respondent level. The red dots with red bars indicate significance at the 5% level, whereas the red triangles with gray bars indicate significance only at the 10% level. The gray dots with gray bars indicate that there is no significance at either conventional level.

C.6 Determinants of Switching to Alternative Options

Figures A10 and A11 isolate the determinants of abstention and blank-ballot choices in Conjoint Study 1 by restricting the comparison solely to active candidate choices versus each expressive non-choice behavior separately. In Figure A10, the outcome compares choosing either candidate (0) with abstaining (1). In Figure A11, the outcome compares choosing either candidate (0) with casting a blank/null vote (1).

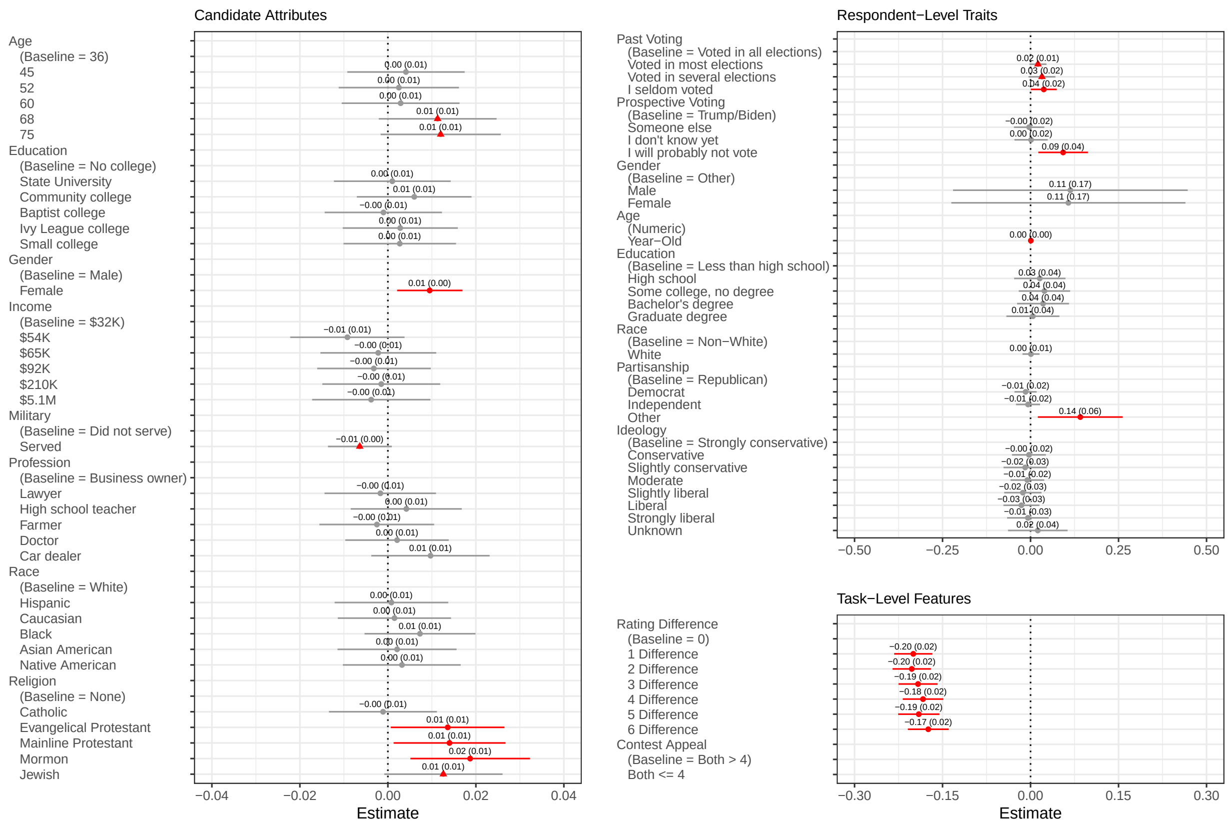

Figure A10: Determinants of Abstention in Conjoint Study 1. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of abstention: candidate choices (Candidate 1 or Candidate 2) = 0; *abstention* = 1. Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

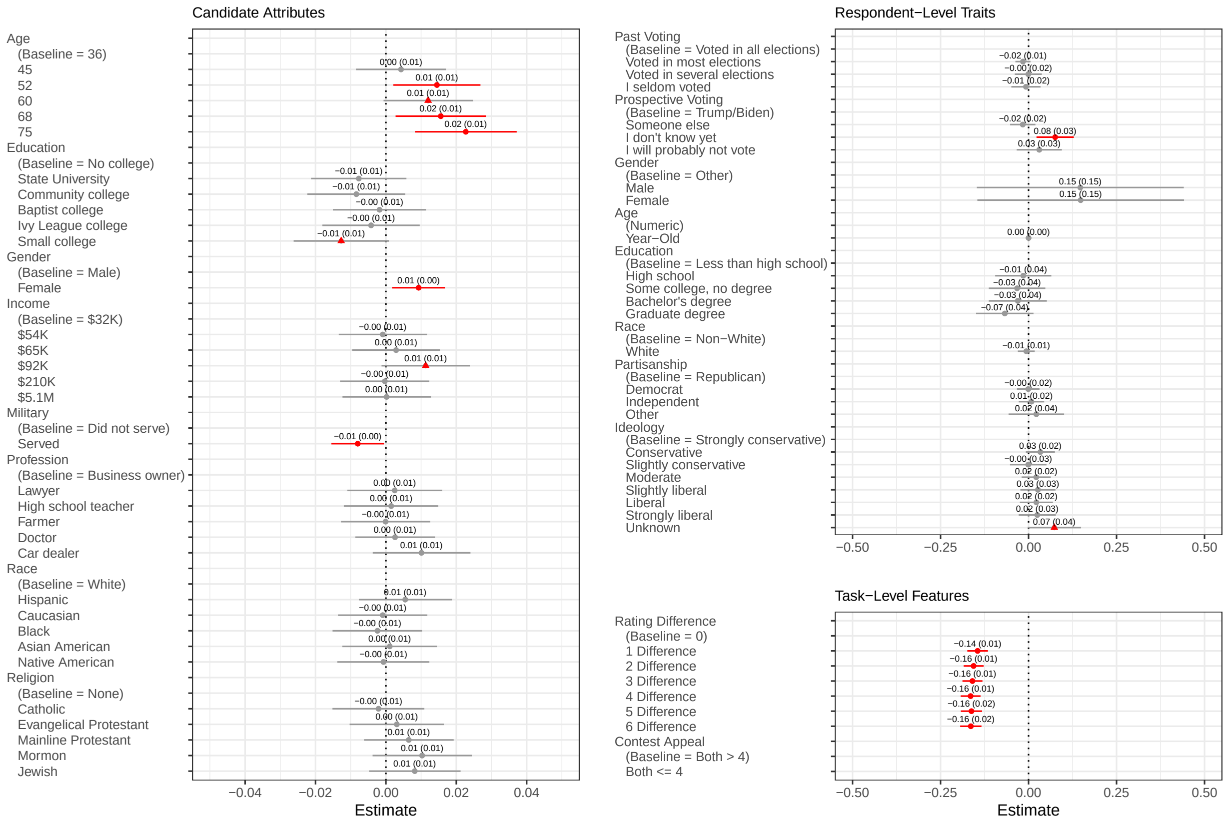

Figure A11: Determinants of Casting a Blank Vote in Conjoint Study 1. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of abstention: candidate choices (Candidate 1 or Candidate 2) = 0; *casting a blank vote* = 1. Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

In Conjoint Study 2, Figures A12 and A13 show that opting out—whether abstaining or casting a blank/null vote—is shaped overwhelmingly by respondent- and task-level characteristics rather than candidate attributes. As in Study 1, respondents with weaker turnout histories and lower intended participation are substantially more likely to opt out, and task conditions that signal low contest quality (small rating differentials or two poorly rated candidates) strongly increase non-choice behavior. Two respondent traits stand out. First, ideology exhibits a clear gradient: relative to strong conservatives, moderates and liberals are consistently less likely to abstain or submit a blank vote. Second, education is positively associated with opting out: respondents with higher levels of education are significantly more likely to withdraw from the choice. These patterns reinforce that opt-out behavior in Study 2 is structured and politically meaningful, not random noise.

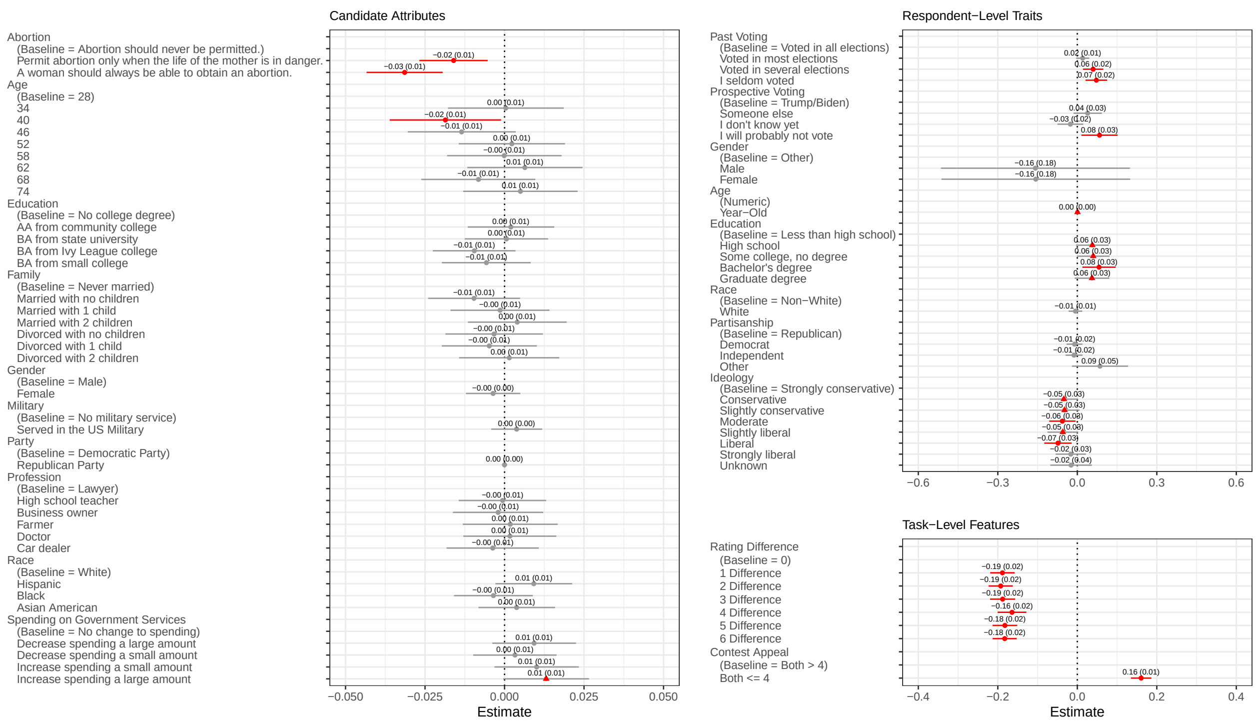

Figure A12: Determinants of Abstention in Conjoint Study 2. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of abstention: candidate choices (Candidate 1, Candidate 2, or blank) = 0; *abstention* = 1. Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

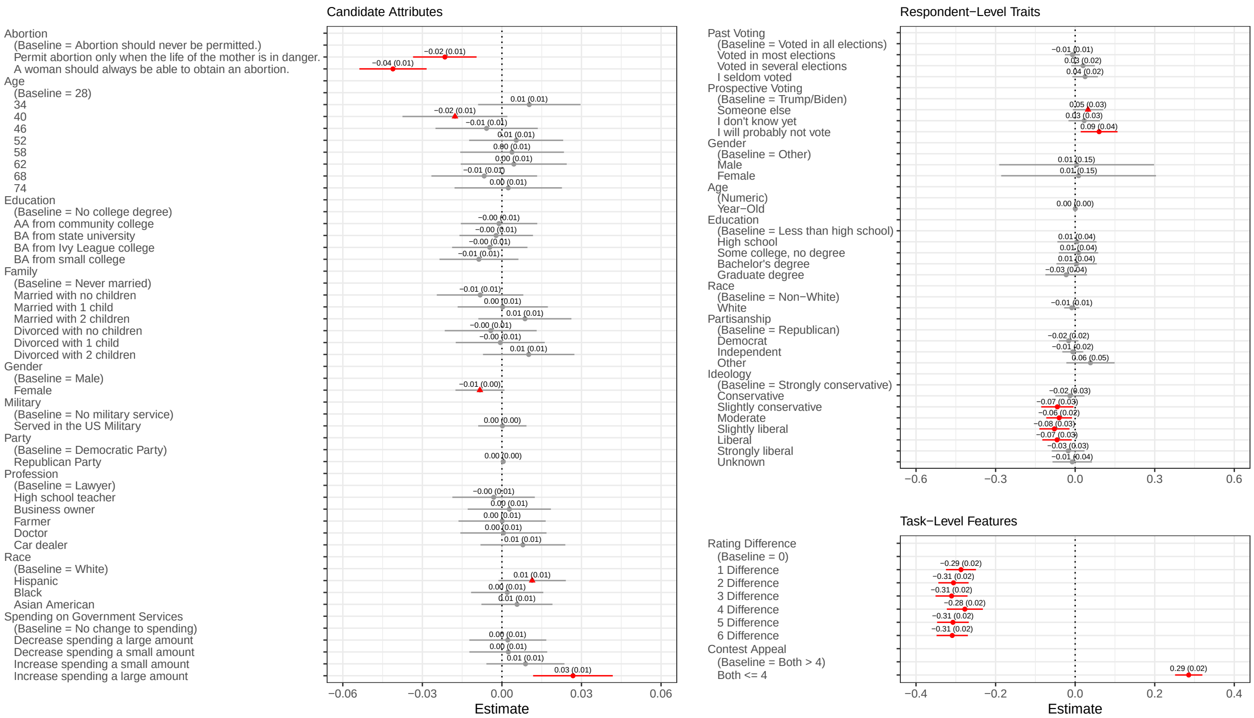

Figure A13: Determinants of Casting a Blank Vote in Conjoint Study 2. Estimates are derived from a linear regression model with respondent-clustered standard errors. The outcome is a binary indicator of abstention: candidate choices (Candidate 1 or Candidate 2) = 0; *casting a blank vote or abstaining* = 1. Red dots mark 5% significance; red triangles mark 10% significance; gray dots indicate no significance. The left panel presents effects of candidate-level attributes, while the right panel shows effects of respondent- and task-level characteristics estimated within the same model.

C.7 Addressing Potential Concerns

One may wonder whether we can fully trust respondents' self-reports of past turnout, given the well-known overreporting in the turnout literature. To mitigate social desirability bias, we used standard "face-saving" wording, reminding respondents that many people cannot or do not vote for perfectly reasonable reasons (registration problems, illness, and lack of time). This is the same strategy commonly used in survey research to reduce turnout inflation. Importantly, if respondents still feel uncomfortable admitting non-voting, we would expect very few to select "seldom vote." Instead, in our sample, "seldom vote" responses outnumber "several times" by a wide margin, suggesting that respondents were comfortable with reporting lower levels of electoral participation. Figure A14 presents the full distribution of self-reported past voting frequency.

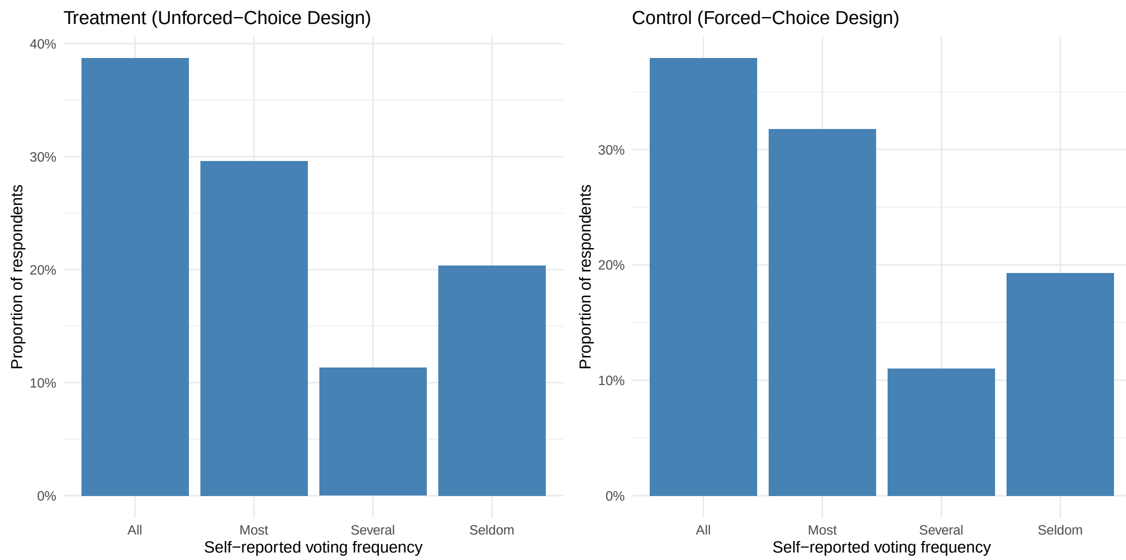

Figure A14: How Often Respondents Say They Have Voted. Bars show the proportion of unique respondents reporting each level of past voting frequency. Proportions are calculated after deduplicating by respondent ID.

One may also be concerned about potential post-treatment bias since the rating-based question was asked after the choice-based question. We acknowledge that this ordering could, in principle, influence task-level features, particularly rating differences. Ideally, we could randomize the order between the choice and rating questions to eliminate such bias. We considered this but still chose not to randomize the question order for two reasons. First, one of our primary objectives is to compare design-induced differences between forced-choice and unforced-choice settings. Randomizing the question order would require doing so in both the treatment and control (forced-choice) designs, which would alter the structure of the canonical conjoint format we seek to benchmark against. Second, even if we randomized the question order, we could not randomize it within respondents across tasks without disrupting respondent attention and consistency. Given these constraints, we decided to maintain a fixed order, which we view as a necessary compromise.

Although we cannot fully rule out post-treatment bias in estimating the effects of task-level features on opt-out decisions, evidence from the control group—where abstention and protest-vote options were not available—suggests that such concerns are limited. More than half of the respondents in the control condition assigned identical ratings to both profiles at least once, indicating that their evaluative judgments were not tightly anchored to the preceding choice-based outcome. If anything, this proportion is a lower bound since some respondents may align their ratings with their selections.

D Pre-Conjoint Survey Questions

D.1 Voting Propensity

A1. Past Voting In asking people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, they didn't have time, or something else happened to prevent them from voting. And sometimes, people who usually vote or who planned to vote forget that something unusual happened on Election Day one year that prevented them from voting that time. So please think carefully for a minute about the recent elections, and other past elections in which you may have voted and answer the following questions about your voting behavior.

How often have you participated in voting since you got the right to vote? Have you voted in all elections, most elections, some elections, or you seldom voted?

- Voted in all elections
- Voted in most elections
- Voted in several elections
- I seldom voted

A2. Future Voting If the 2024 presidential election were between Donald Trump for the Republicans and Joe Biden for the Democrats, would you vote for Donald Trump, Joe Biden, someone else, probably not vote, or do not know yet?

- Donald Trump
- Joe Biden
- Someone else
- I do not know yet what candidate I will support
- I will probably not vote

B. Personal Background

B1. Gender What is your gender?

- Man
- Woman
- Other

B2. Age How old are you (in years)? ____________________

B3. Education What is the highest degree that you have earned?

- Less than high school credential
- High school diploma
- Some post-high school, no bachelor's degree
- Bachelor's degree
- Graduate degree

B4. Race/Ethnicity What racial or ethnic group or groups best describe you?

- Black
- Asian
- Native American
- Hispanic or Latino
- White
- Other: ___________

B5. Partisanship In politics today, do you consider yourself as a Republican, a Democrat, an Independent, or what?

- Republican
- Democrat
- Independent
- Other: ___________

B6. Ideology We hear a lot of talk these days about liberals and conservatives. On a seven-point scale on which the political views that people might hold are arranged from extremely liberal (1) to extremely conservative (7), where would you place YOURSELF on this scale, or haven't you thought much about this?

- Strongly conservative (1)
- Conservative (2)
- Slightly conservative (3)
- Moderate (4)
- Slightly liberal (5)
- Liberal (6)
- Strongly liberal (7)
- I don't know