Marginal Component Effects from Conjoint Experiments Have Predictive Validity: Evidence from Individual-Level Analyses

Kirill Zhirkov[1], Fabian Neuner[2], and Melle Scholten[3]

[1] University of Virginia

[2] Arizona State University

[3] University of Twente

[Draft. Please don't cite or circulate]

Abstract

Marginal component effects (MCEs) are the most common quantities used to present results from conjoint survey experiments, but recent scholarship has questioned their appropriateness as measures of preferences. Much of this debate concerns the accuracy of the signs and magnitudes of MCEs, which is assessed through aggregate-level analyses. In this paper, we explore the predictive validity of MCEs—the degree to which they map onto measures of other relevant concepts—through individual-level analyses. Our evidence comes from three studies, two in the United States and one in Germany, in which respondents both choose between and rate hypothetical candidates in conjoint experiments—and then report their support for existing parties and policies in separate questions. Using these data, we show that respondents' preferences measured via individual MCEs in conjoint experiments strongly and consistently predict self-reported party support, voting intentions, and issue positions in the expected directions.

Keywords: conjoint experiments, marginal component effects, predictive validity, preferences

1 Introduction

Conjoint experiments are a class of survey experiments designed to study multidimensional choices (Bansak, Hainmueller, et al. 2021). In a conjoint task, respondents are presented with hypothetical options described in terms of several attributes with randomized values and asked to rate these options, choose among them, or both. The main benefit of the conjoint-experimental design is that it allows researchers to independently estimate and compare the effects of multiple considerations or factors on people's choices. Conjoint tasks also have other beneficial properties: lower susceptibility to satisficing (Bansak et al. 2018), stable performance when unlikely attribute combinations are included (Bansak and Jenke 2025), and the capacity to partially alleviate social desirability concerns (Horiuchi et al. 2022). The power and flexibility of conjoint experiments have thus quickly made them a popular tool in the empirical study of politics.[1]

The most common quantities used to present results from conjoint experiments are known as marginal component effects (MCEs): differences in choice probabilities or rating scores if attribute values change. Originally, these quantities were introduced to political science as causal estimands conceptually similar to treatment effects in standard survey experiments (Hainmueller et al. 2014). However, the practical application of conjoint experiments—and thus MCEs—in the literature primarily concerns the measurement of preferences. For instance, a widely cited conjoint study has found that Americans prefer high-skilled immigrants over low-skilled ones and that these preferences are stronger than those regarding immigrants' countries of origin (Hainmueller and Hopkins 2015).

While conjoint experiments' ability to provide causal estimates is a point of consensus among political methodologists, the use of MCEs as measures of preferences is increasingly contested. The issues raised in the literature regarding MCEs include their

[1] We identified at least 59 papers published between 2015 and 2024 in the three leading general-interest political science journals—the *American Political Science Review*, *American Journal of Political Science*, and the *Journal of Politics*—that used conjoint experiments, of which 15 were published in 2024 alone.

aggregation (Abramson et al. 2022), reliability (Clayton et al. 2023), external validity (de la Cuesta et al. 2022), and interpretation (Leeper et al. 2020). Another line of methodological work addresses differences between choice and rating outcomes in conjoint designs as well as some limitations of the former (Ganter 2023; Miller and Ziegler 2024; Treger 2025). These criticisms are important, but they exclusively concern how accurately and reliably MCEs estimate and represent population preferences. The evidence corroborating the hypothesis that MCEs correctly measure preferences in terms of sign and magnitude also almost exclusively comes from aggregate-level results (Bansak et al. 2022; Hainmueller et al. 2015; Jenke et al. 2021).

In this paper, we advance this methodological debate by introducing a novel way to explore the validity of conjoint MCEs. Instead of assessing the accuracy of specific population-level estimates, we focus on how well individual-level MCEs map onto the measures of other theoretically relevant concepts—that is, on their *predictive validity*. For instance, do people who prefer candidates promising to deport immigrants in conjoint experiments also express support for existing parties that run on strict immigration enforcement? In other words, we shift attention from whether conjoint MCEs have *correct values* to whether they capture *meaningful variation* in respondents' preferences. Given the hypothetical nature of conjoint tasks, this question is important—but it remains unanswered due to the absence of individual-level evidence.

Our test focuses on one of the most popular applications of conjoint experiments in political science: measuring preferences toward the attributes of political candidates (Carnes and Lupu 2016; Costa 2021; Hanretty et al. 2020). Our data come from three studies, two in the United States and one in Germany. We use conjoint tasks to estimate individual MCEs, which provide respondents' preferences regarding candidates' demographic traits and issue positions (Zhirkov 2022). We then show that issue preferences measured through conjoint experiments strongly and consistently predict respondents' party affect, vote intentions, and policy opinions in the expected directions. We also estimate individual MCEs using both rating and choice outcomes and demonstrate that they are equally good predictors of the relevant individual-level covariates. These results make an important contribution to the ongoing debate about the measurement properties of conjoint MCEs by showing that, independently of the outcome format, they capture meaningful variation in respondents' preferences.

2 The Estimation and Interpretation of Conjoint MCEs

The MCE is the most popular causal estimand used in the analysis of conjoint experiments. It represents the (marginal) effect of a single attribute value (component) on the choice probability or rating score against the baseline value of the same attribute. For instance, how would preference for a hypothetical candidate change if they were a woman rather than a man? MCEs can be defined at the population or aggregate level (Hainmueller et al. 2014), at the level of individual respondents (Zhirkov 2022), or even at the level of specific observations or profiles (Robinson and Duch 2024).

Practical applications of conjoint analysis in political science are mainly concerned with measuring preferences. Examples include conjoint experiments on candidate choice (Carnes and Lupu 2016; Costa 2021; Hanretty et al. 2020), policy proposals (Ballard-Rosa et al. 2017; Bansak, Bechtel, et al. 2021), immigrant admission (Hainmueller and Hopkins 2015), news sources (Mummolo 2016), residential choices (Mummolo and Nall 2017), and many others. In such applications, conjoint MCEs are interpreted as measures of both respondents' preferences regarding attribute values (i.e., which attribute values make options more attractive) and the relative importance of different attributes (i.e., which attributes play larger roles in respondents' choices). For instance, an MCE for the "College" value of the education attribute in a candidate-choice conjoint experiment that is positive and reliably different from zero would be interpreted as evidence that, first, voters prefer college-educated politicians and, second, that a candidates' education level is an important factor driving voters' choices.

Consider this formal illustration that concerns the version of the MCE most widely used in the literature: the AMCE, which averages the effect of interest across observations. In practice, AMCEs are estimated using OLS regression models with categorical predictors corresponding to attribute values (with one value serving as the baseline). In a conjoint experiment, respondents indexed $i \in \{1, \ldots, I\}$ are asked to rate profiles of hypothetical candidates indexed $j \in \{1, \ldots, J\}$. Let a candidate's position on immigration be one of the attributes with two possible values, "Admit" and "Deport," randomized independently from other attributes. Then, the AMCE can be estimated using the following equation:

$$\text{rating}_{ij} = \alpha + \beta[\text{Immigrants} = \text{Deport}]_{ij} + \varepsilon_{ij}, \tag{1}$$

where α is the expected rating score for a hypothetical candidate who would admit immigrants (the baseline) and β is equivalent to the AMCE of interest. In this case, the AMCE is the expected difference in rating scores between the baseline and a candidate who would deport immigrants. A positive AMCE means that voters, on average, prefer candidates who would deport immigrants over those who would admit them, and vice versa. The magnitude of the AMCE, in turn, represents the intensity of that preference in the population. In other words, the AMCE—or any conjoint MCE for that matter—is interpreted as both a causal quantity and a measure of preferences.

3 The Validity of Conjoint MCEs as Measures of Preferences

While the causal interpretation of MCEs is universally accepted, their usage as measures of preferences is currently a topic of debate in the methodological literature. On the one hand, researchers have identified multiple sets of plausible conditions when MCEs can return estimates that are incorrect in sign or magnitude. This can happen when measurement error is too high (Clayton et al. 2023), the real-world distributions of attributes differ from the ones assumed by the analyst (de la Cuesta et al. 2022), there are profile pairs for which an attribute has the same value (Ganter 2023), or respondents are not able to "abstain" from making a choice (Miller and Ziegler 2024). All these contributions concern the accuracy of specific AMCE estimates: whether they correctly capture the signs and magnitudes of the corresponding population-level preferences. Nevertheless, the studies cited above suggest that MCEs as measures can be subject to both systematic bias and random error—as well as that conjoint designs with forced-choice outcomes are particularly prone to these problems.

On the other hand, contrary to these criticisms, there are studies that defend the use of conjoint MCEs as measures of preferences. For instance, there exists a formal argument in favor of measures that, like conjoint AMCEs, incorporate preference intensity (Bansak et al. 2022). Empirically, AMCE estimates obtained from conjoint experiments closely approximate real-world behavior, such as observed vote results in Swiss naturalization referendums (Hainmueller et al. 2015). Furthermore, attribute importance revealed in conjoint experiments corresponds to respondents' attention to the same attributes assessed via eye-tracking (Jenke et al. 2021). But this evidence, again, focuses on the accuracy of specific estimates, comes from the aggregate level, and does not shed light on how MCEs map onto relevant *attitudinal* covariates.

Much of the current debate concerning conjoint experiments thus ultimately boils down to the question of *measurement validity*, or whether MCE scores meaningfully capture preferences, which is the concept they are intended to measure. Thus far, the methodological literature has addressed the extent to which MCEs accurately recover the signs and magnitudes of specific preferences. But this question is difficult to answer definitively, because respondents' true preferences are fundamentally unobservable.

Here, we extend the debate regarding the measurement validity of conjoint experiments and address an important aspect of it, which has been largely ignored in the literature to date: do MCEs capture meaningful variation in preferences? In other words, we answer the question of nomological or *predictive* validity: whether the measure of interest replicates well-studied and theoretically motivated relationships with measures of other concepts (Adcock and Collier 2001). To achieve this goal, we estimate conjoint MCEs for individual respondents and assess how well they predict relevant self-reported political attitudes, behavioral intentions, and policy opinions.

Exploring how well conjoint MCEs map onto measures of other concepts of interest can make several important methodological contributions. First, as of now, there is no evidence pertaining to the predictive validity of conjoint MCEs at the individual level. To show the latter, respondents' preferences revealed via hypothetical choices made in conjoint experiments (e.g., a position on a political issue) should predict relevant self-reported attitudes and intentions (e.g., planning to vote for a party with the same issue position). Second, if conjoint-estimated preferences can consistently predict the covariates of interest, it would also mean that those estimates are reliable enough to be of use to applied researchers. Finally, it remains unclear whether using forced-choice as opposed to rating outcomes affects the extent to which preferences revealed in conjoint experiments correspond with self-reported attitudes toward relevant political parties and issue positions.

To summarize, the methodological literature focuses on the accuracy of specific MCE *values* while largely ignoring whether conjoint MCEs capture meaningful *variation* in preferences. As a result, the predictive validity of conjoint MCEs—that is, the extent to which they map onto theoretically relevant covariates—has not been demonstrated to date. There is also no evidence on how this validity may differ across the two common outcomes in conjoint experiments: interval ratings and forced choices. Ultimately, these questions can only be answered when MCEs are estimated, and their predictive power is assessed, at the individual level.

4 Individual-Level MCE Estimation

Fortunately, respondent-level preferences in conjoint experiments can be estimated in the form of individual MCEs or IMCEs (Zhirkov 2022). Their estimation does not require any additional assumptions compared to AMCEs and uses the same method: OLS regression. Consider the example introduced previously with a candidate conjoint experiment and position on immigrants as the attribute. Assume that each respondent is asked to rate multiple profiles, which is the standard in modern conjoint tasks. To obtain IMCEs, one should simply take the same expression as presented in Equation 1 and estimate it independently for each respondent:

$$\text{rating}_{ij} = \alpha_i + \beta_i[\text{Immigrants} = \text{Deport}]_{ij} + \varepsilon_{ij}. \tag{2}$$

In this equation, α_i is the individual (respondent-specific) expected rating score for a candidate who would admit immigrants, and β_i is equivalent to the IMCE of interest. The sign and magnitude have the same interpretations as those of the AMCE: in this context, it measures the direction and strength of a relative preference for a candidate who would deport immigrants over one who would admit them—but the IMCE can be estimated for each individual respondent in the analyzed sample.

Originally, IMCEs were proposed as a method to explore the heterogeneity of preferences in the population that does not rely on apriori categorizations. However, they can easily be used to test predictive validity of conjoint MCEs at the individual level. In survey studies that include both the conjoint component (e.g., a hypothetical candidate choice task) and questions about relevant attitudes and behaviors (e.g., support for existing political parties), IMCEs can be used to predict the political covariates of interest. This setup shares similarities with the method of using predicted responses from list experiments as explanatory variables in regression models (Imai et al. 2015).

Consider the estimate of the IMCE for the value “Deport” of the attribute “Position on immigrants” in the candidate conjoint experiment from Equation 2. It reflects the direction and intensity of an individual respondent’s preference for candidates who would deport immigrants over those who would admit them (higher positive values of the IMCE indicate stronger preference for “Deport,” and vice versa). Assume that the survey data also includes a question on respondents’ support for an existing anti-immigration party. Then, IMCE estimates can be used as predictors of this support in a regression:

$$\text{support}_i = \gamma + \delta(\text{IMCE: Immigrants, deport})_i + u_i. \qquad (3)$$

Parameter γ is the intercept, or the expected support for an anti-immigration party for a voter with an IMCE of zero (i.e., perfect indifference between admitting and deporting immigrants). Parameter δ represents the estimated association between preference for hypothetical candidates who would deport immigrants in a conjoint experiment and self-reported support for an existing anti-immigration party.[2]

5 Analytic Approach

The estimation of IMCEs allows us to conduct novel individual-level analyses that can shed light on how well conjoint MCEs capture variation in respondents’ preferences. First, we test whether IMCEs exhibit predictive validity at the individual level. To do that, we estimate regression models, in which IMCEs are used to predict self-reported party support, voting intentions, and issue preferences. Coefficients with expected signs that are reliably different from zero would indicate predictive validity of conjoint IMCEs.

Second, we test whether conjoint IMCEs estimated using rating and choice outcomes perform equally well. To do that, we compare the predictive power of IMCEs estimated using rating and choice outcomes in regression models. Equal predictive power assessed via the coefficient of determination statistic would indicate that conjoint IMCEs estimated

[2] IMCE estimates are based on relatively small numbers of observations, and we run robustness checks to account for the resulting uncertainty (Zhirkov 2022). See Section A of the SI for the estimation details.

using rating and choice outcomes have the same predictive validity. We implement these tests in three different studies, two in the United States and one in Germany.

6 Study 1

6.1 Data

We fielded an original online survey with a conjoint-experimental component in June 2024. Respondents were recruited using Cint Theorem, a popular source of convenience samples with demographics mirroring national benchmarks (formerly Lucid; Coppock and McClellan 2019). A total of 825 respondents completed the survey.[3]

In the conjoint-experimental task, respondents were presented with paired profiles featuring hypothetical politicians described as "potential House candidates." The number of unique profiles per respondent was 16. Respondents were asked to both indicate which of the candidates from each pair they preferred (forced choice) and to rate each candidate on a scale from 0 = *Definitely would not consider voting for* to 10 = *Definitely would consider voting for*. The candidates were described in terms of six attributes with randomized values: three demographic traits (age, gender, and race) and three issue positions (size of government, abortion, and unauthorized immigrants). Demographic traits were always presented before issue positions in the conjoint table. The order of the attributes within these two subgroups was randomized between respondents. See Table 1 for potential values of the six attributes. All attribute values were independently randomized with uniform distributions. See Section B of the Supplementary Information (SI) for an example of candidate profiles as presented to respondents.

After the conjoint task, respondents were asked about their party support, vote intention in the upcoming general election, and their stances on various issues. Affect toward the Democratic Party and the Republican Party was measured using standard feeling thermometer scales ranging from 0 = *Very cold* to 100 = *Very warm*. We included

[3] 21 respondents showed no variation in profile rating scores, so IMCEs for them could not be estimated.

Table 1. Candidates' attributes, Study 1

Attribute	Values
Age	*Younger:* 30–49
	Older: 50–69
Gender	Male
	Female
Race	White
	Black
	Hispanic
	Asian
Size of government	Government should provide more services and raise taxes
	Government should provide fewer services and cut taxes
Abortion	A woman should be able to obtain an abortion
	Abortion should not be permitted
Unauthorized immigrants	Allow unauthorized immigrants to remain in the United States
	Send unauthorized immigrants back to their home countries

both a question on respondents' 2024 presidential vote intention (Biden vs. Trump) and a generic congressional ballot question (Democrat vs. Republican).[4] Self-reported issue positions corresponded to the ones included in the conjoint experiment: size of government, abortion, and unauthorized immigrants. See Section C of the SI for exact questions and response options. Respondents' demographic information was provided by Cint.

6.2 Tests

We predict party affect and vote intentions with IMCE-measured preferences regarding candidates' demographics and issue positions. Given the ideological positioning of the two major U.S. parties and their presidential candidates in the 2024 election, we interpret the following associations as evidence in favor of predictive validity:

- Conjoint-estimated preferences for smaller (as opposed to bigger) government should be positively associated with affect toward the Republican Party, intentions to vote for Donald Trump and a Republican congressional candidate, and self-reported

[4] When we fielded the survey, Joe Biden was still the Democratic presidential nominee.

preference for smaller government; they should be negatively associated with affect toward the Democratic Party.

- Conjoint-estimated preferences for abortion to be illegal (as opposed to legal) should be positively associated with affect toward the Republican Party, intentions to vote for Donald Trump and a Republican congressional candidate, and self-reported preference for abortion to be illegal; they should be negatively associated with affect toward the Democratic Party.
- Conjoint-estimated preferences for deporting unauthorized immigrants (as opposed to allowing unauthorized immigrants to remain in the United States) should be positively associated with affect toward the Republican Party, intentions to vote for Donald Trump and a Republican congressional candidate, and self-reported preferences for deporting unauthorized immigrants; they should be negatively associated with affect toward the Democratic Party.

6.3 Results

As outlined above, we estimate IMCEs from the conjoint experiment and then use them to predict Democratic and Republican feeling thermometers as well as presidential and congressional vote intentions.[5] Theoretically, estimated IMCEs can range from -10 (a respondent always rates profiles having the corresponding attribute value at 0 and profiles not having the value at 10) to 10 (a respondent always rates profiles having the corresponding attribute value at 10 and profiles not having the value at 0). Empirically, 99% of estimated IMCEs are approximately between -6 and 6 (see Section E of the SI for the empirical densities).

[5] We also estimate AMCEs by respondent's party. The results are presented in Section D of the SI. They show that Democrats strongly prefer candidates with liberal positions on abortion and, to a lesser extent, size of government and unauthorized immigrants. Republicans, in turn, prefer candidates with conservative positions on immigration—and it is the only issue position with a significant AMCE.

The results are presented in Figure 1 as OLS regression coefficients (Democratic and Republican feeling thermometers) and logistic regression coefficients (vote intentions). They show that conjoint-estimated preferences regarding candidates' demographic traits—age, gender, and race—are mostly inconsequential. Preferences regarding the size of government show relatively weak associations with both party affect and vote intentions. Preferences on abortion and immigration, however, produce sizable coefficients that are reliably different from zero and have expected signs: respondents preferring candidates who believe abortion should be illegal and favor deporting unauthorized immigrants feel warmer toward the Republican Party and colder toward the Democratic Party. IMCE-measured conservative issue preferences also predict intent to vote both for Trump in the presidential election and for a Republican in the generic congressional ballot.

Figure 2 further demonstrates that the associations between preferences measured through the conjoint experiment and self-reported support for the existing parties are substantial in terms of magnitude. The figure uses preferences on the immigration issue and party affect to illustrate this. A respondent with the strongest observed IMCE-measured preference against deporting unauthorized immigrants (or in favor of letting them stay in the United States) is predicted to report very cold feelings toward the Republican Party (15 degrees on the 0–100 scale) and very warm feelings toward the Democratic Party (90 degrees). For those with the strongest observed preference in favor of deporting unauthorized immigrants (or against letting them stay) the picture is the opposite: the regression model predicts warm feelings (about 80 degrees) toward the Republican Party and cold feelings toward the Democratic Party (around 15 degrees).

We also examine predictive validity by exploring how well conjoint-estimated positions on political issues predict self-reported positions on the same issues. The results are presented in Figure 3 as predicted probabilities of taking a conservative position on a self-reported question depending on the corresponding conjoint IMCE. They show an

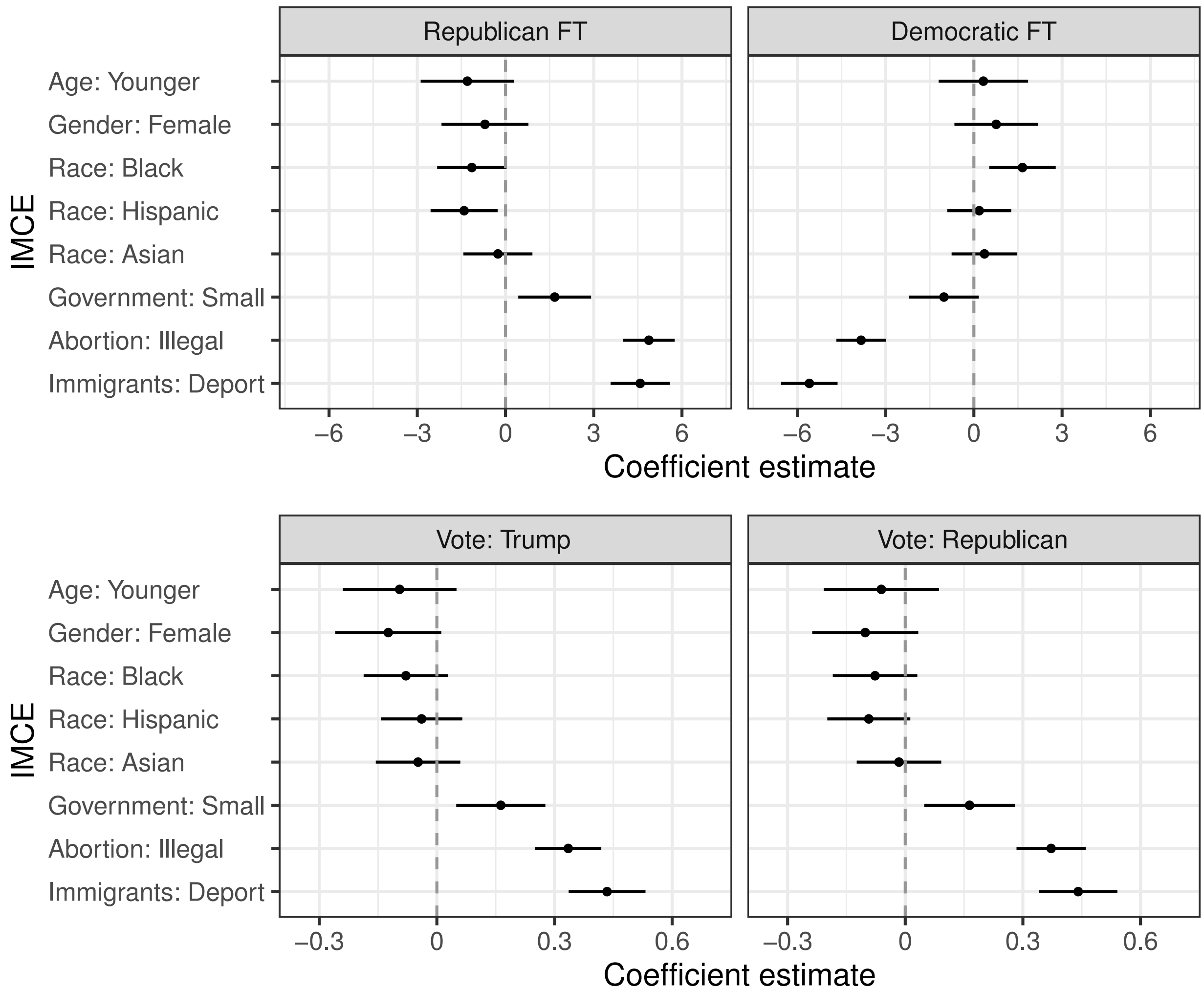


Figure 1. Conjoint IMCEs as predictors of party affect (OLS regressions) and vote intentions (logistic regressions), Study 1
Note. Point estimates with 95% confidence intervals. See Section F of the SI for uncertainty-adjusted estimates.

extremely close correspondence between the conjoint measures and the self-reported measures of preferences on all three issues. Overall, our results show that IMCEs exhibit strong predictive validity.

Another important methodological question concerns the limitations of using forced-choice outcomes in conjoint experiments as opposed to potentially more informative

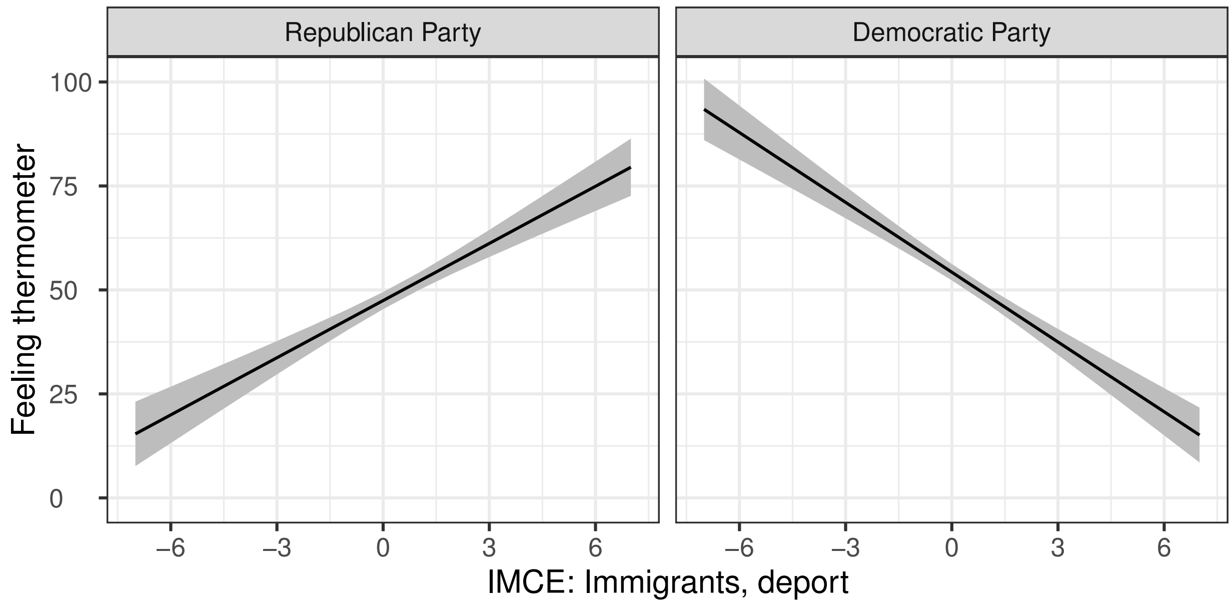


Figure 2. Predicted affect toward the Republican Party and the Democratic Party depending on IMCE-measured immigration policy preference, Study 1
Note. Shaded areas are 95% confidence intervals.

interval ratings. We examine this question by once again turning to predictive validity and exploring how well IMCEs calculated from rating and choice outcomes predict individual covariates of interest. To implement the comparison, we estimate IMCEs for each of the three issue attributes—size of government, abortion, and unauthorized immigrants—using forced-choice outcomes, in addition to IMCEs based on rating outcomes. The procedure is similar: it uses respondent-specific OLS regressions to obtain IMCE estimates.

Then, we use rating- and choice-based IMCEs to predict the same covariates: party affect and vote intentions. The results are presented in Figure 4 using R^2 (for OLS regressions) and McFadden's pseudo-R^2 (for logistic regressions) statistics with bootstrapped 95% confidence intervals as the measures of prediction quality. This shows that IMCEs based on rating and choice outcomes exhibit no differences in prediction quality. Practically, this result confirms that individual MCEs have predictive validity, even when underlying conjoint designs use forced-choice outcomes.

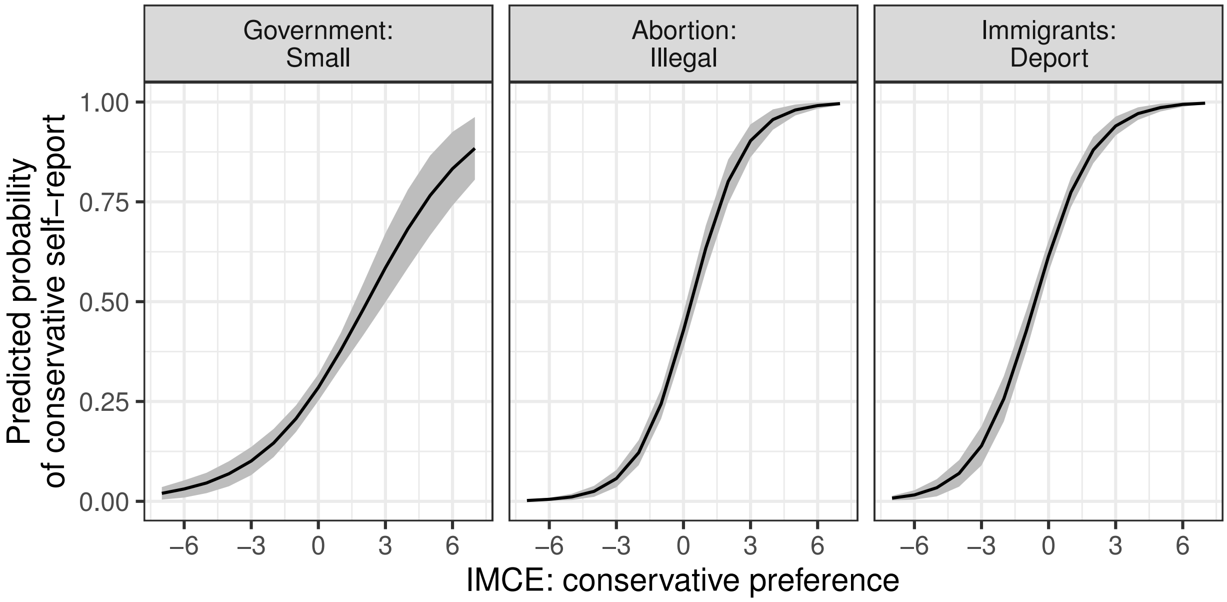


Figure 3. Predicted probabilities of taking conservative positions on the three issues depending on IMCE-measured conservative preferences, Study 1
Note. Shaded areas are 95% confidence intervals.

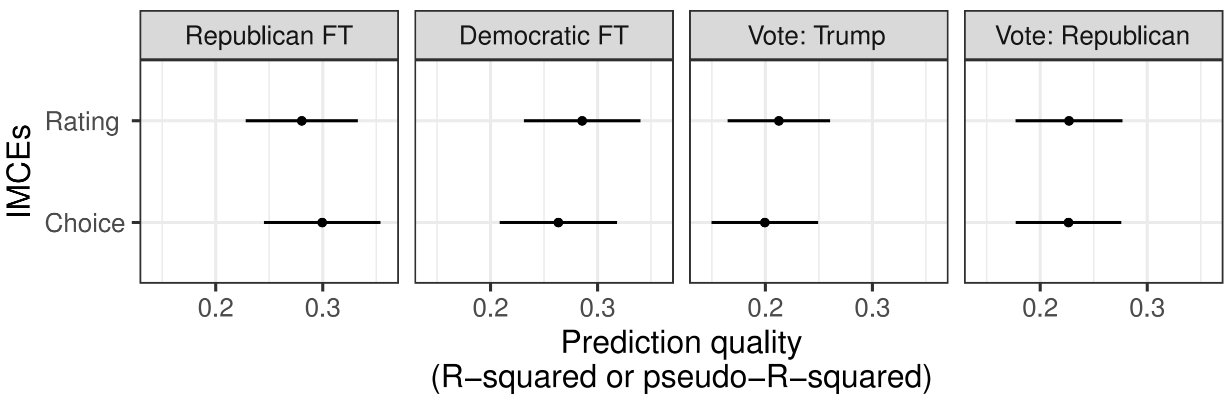


Figure 4. Prediction quality for party affect and vote intentions depending on IMCE estimation method, Study 1
Note. Point estimates with bootstrapped 95% confidence intervals.

7 Study 2

To examine the robustness of our results in the United States, we replicated them in a separate study. The data was collected on CloudResearch, another popular and reputable source of convenience samples (Chandler et al. 2019). Design-wise this was an exact replication of Study 1, and the final sample was 839.

We use these data to replicate the same validity tests: use IMCEs to predict self-reported support for existing political parties, and then assess predictive power of rating- and choice-based IMCEs. The tests return results exactly in line with those obtained in Study 1 and confirm predictive validity of conjoint MCEs. First, IMCEs strongly and consistently predict party affect and vote intentions (see Section G of the SI). Second, IMCEs based on rating and choice outcomes exhibit equal prediction quality (see Section H of the SI).

8 Study 3

8.1 Data

To demonstrate that our results are generalizable beyond the U.S. two-party context, we implement an additional predictive validity test using data from a multiparty system. In Study 3, we leverage an online survey with a conjoint-experimental component fielded on a national sample in Germany in 2017 (Neuner and Wratil 2022). The conjoint-experimental part of the survey presented respondents with ten profiles of hypothetical candidates in five pairs. Respondents were asked to make choices within pairs and also rate each profile on a scale from 1 = *Cannot imagine voting for this candidate* to 7 = *Can easily imagine voting for this candidate*. Candidates were described in terms of their positions on four issues: two left–right issues (taxing the rich and refugee admission) and two mainstream–anti-establishment issues (globalization and the European Union). Each issue attribute had four potential values, but we collapsed them into two values per issue to be able to estimate IMCEs with only ten profiles per respondent. When collapsing, we preserved the direction of preferences. For instance, we collapsed positions "For the admission of some new refugees" and "For the admission of a great many new refugees" ("Admit") versus "For the deportation of some refugees" and "For the deportation of a

Table 2. Candidates' issue positions, Study 3

Issue	Collapsed	Original
Taxes on the rich	Higher	Much higher taxes on the rich
		Somewhat higher taxes on the rich
	Lower	Somewhat lower taxes on the rich
		Much lower taxes on the rich
Refugees	Admit	Admit a great many new refugees
		Admit some new refugees
	Deport	Deport some refugees
		Deport a great many refugees
Free trade	More	Much more globalization
		Somewhat more globalization
	Less	Much less globalization
		Somewhat less globalization
European Union	Integrate	Develop the EU into a common state
		Stronger cooperation within the EU
	Withdraw	Weaker cooperation within the EU
		Germany's withdrawal from the EU

great many refugees" ("Deport"). Issue positions and the way they were coded are presented in Table 2. Similar to the U.S. studies, all attribute values were independently randomized with uniform distributions.

In addition to completing the conjoint task, each respondent was asked which party they would vote for if the next election happened the following Sunday. Here, we focus on voting for four larger German parties. Table 3 lists those parties' German abbreviations, English translations of their names, and their positioning on both left vs. right and mainstream vs. anti-establishment dimensions. The sample included 1,640 respondents.

8.2 Tests

We use policy preferences measured via conjoint IMCEs to predict respondents' self-reported vote intentions. Given German parties' positions, the following results can be interpreted as evidence in favor of predictive validity:

Table 3. Parties included in the analysis, Study 3

	Title	Family
AfD	Alternative for Germany	Right, anti-establishment
CDU/CSU	Christian Democratic Union / Christian Social Union	Right, mainstream
Die Linke	The Left	Left, anti-establishment
SPD	Social Democratic Party of Germany	Left, mainstream

- Conjoint-estimated preferences for lower taxes and for deporting refugees should be positively associated with the probability of planning to vote for AfD and CDU (right-wing and center-right parties) and negatively associated with the probability of planning to vote for Die Linke and SPD (left-wing and center-left parties).
- Conjoint-estimated preferences for less globalization and for withdrawal from the European Union should be positively associated with the probability of voting for AfD and Die Linke (anti-establishment parties) and negatively associated with the probability of voting for CDU and SPD (mainstream parties).

8.3 Results

Because of the categorical nature of the dependent variable (intended vote in a multiparty system), we use multinomial logistic regression to predict respondents' vote intentions with IMCE-estimated issue preferences. Theoretically, estimated IMCEs can range from -6 (a respondent always rates profiles having the corresponding attribute value at 1 and profiles not having the value at 7) to 6 (a respondent always rates profiles having the corresponding attribute value at 7 and profiles not having the value at 1). Empirically, 99% of estimated IMCEs are approximately between -4 and 4 (see Section I of the SI for the empirical densities).

Results are presented in Figure 5 as multinomial logistic regression coefficients with the SPD, the party with most intended votes in the data, serving as the baseline choice. They show coefficients that are reliably different from zero and have expected directions.

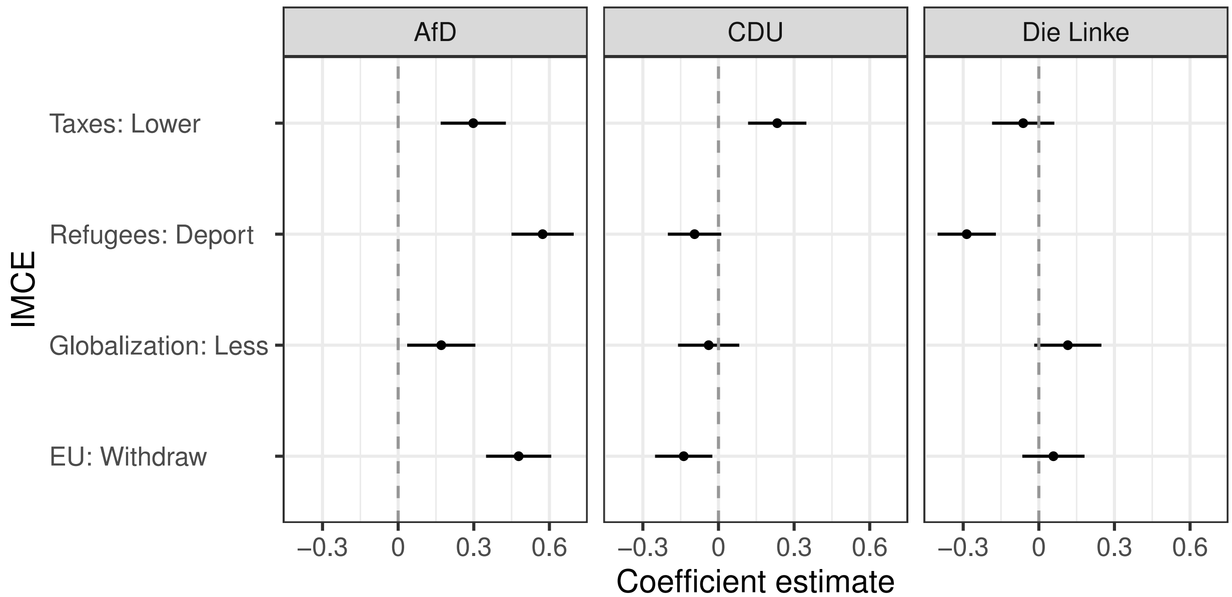


Figure 5. Conjoint IMCEs as predictors of vote intention: coefficients from the multinomial logistic regression, Study 3
Note. Voting for SPD is the baseline category. Point estimates with 95% confidence intervals. See Section J of the SI for uncertainty-adjusted estimates.

When compared to the center-left SPD, respondents who prefer to deport refugees are more likely to vote for the right-wing AfD and less likely to vote for the left-wing Die Linke. Those who prefer lower taxes for the rich are also more likely to vote for the right-wing and center-right parties (AfD and CDU) than for center-left SPD. Finally, opposition to the EU membership and, to a lesser extent, globalization increase the probability of voting for the AfD vis-a-vis the SPD.

Since coefficients from multinomial logistic regressions are difficult to interpret directly, we also present the results in terms of the key quantity of interest: predicted probabilities. Figure 6 shows how IMCE-measured preferences for deporting refugees predicts voting for right-wing AfD and left-wing Die Linke. Specifically, a person with the strongest preferences against deportation (or in favor of admission) is predicted to have an almost zero percent chance of voting for AfD and approximately a 50% chance of voting for

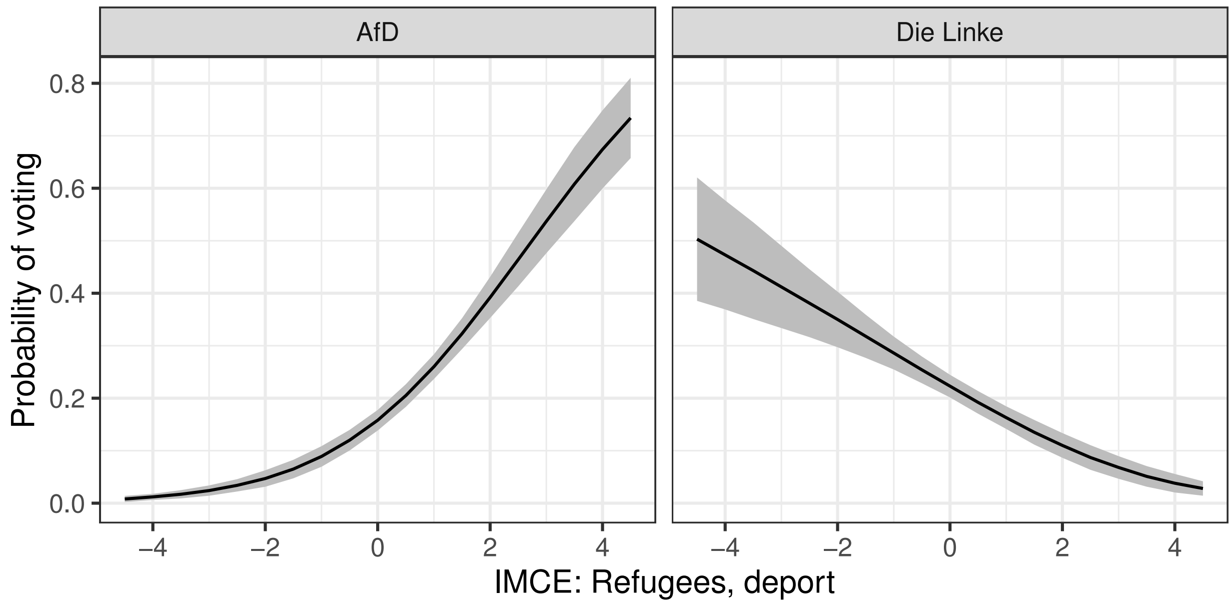


Figure 6. Predicted probabilities of voting for Die Linke and AfD by IMCE-measured refugee policy preference, Study 3
Note. Point estimates with 95% confidence intervals.

Die Linke. For those with the strongest preference for deportation (or against admission), the picture is the opposite: they are predicted to have less than a 5% chance of voting for Die Linke and more than a 70% chance of voting for AfD. Taken together, these findings strongly corroborate the predictive validity of conjoint IMCEs—similar to Studies 1 and 2.

Then, we compare how IMCEs estimated using rating and choice outcomes perform in the same multinomial logistic regression as predictors of vote intention. The results are exactly the same: the corresponding pseudo-R^2 statistics are .08 (95% CI from .07 to .10) for IMCEs based on both rating and choice outcomes. Once again, in line with Studies 1 and 2, we show that rating- and choice-based IMCEs have equivalent predictive power.

9 Discussion and Conclusion

In this paper, we contribute to the ongoing debate in the methodological literature on MCEs, the most popular causal estimands from conjoint survey experiments. While the suitability of MCEs, especially in forced-choice conjoint designs, as measures of preferences has been questioned (Clayton et al. 2023; de la Cuesta et al. 2022; Ganter 2023), there are also formal arguments and empirical evidence in their favor (Bansak et al. 2022; Hainmueller et al. 2015; Jenke et al. 2021). However, this prior work considers the accuracy of specific MCE values but does not engage with the question of whether MCEs capture meaningful variation in preferences. Since conjoint studies are almost exclusively analyzed in the aggregate, they cannot explore how well MCEs map onto measures of other relevant concepts. We address this gap by estimating individual MCEs and exploring their predictive validity. Our data come from three candidate-choice experiments, two in the United States and one in Germany.

Across all three studies, our results show that conjoint MCEs as measures of preferences exhibit predictive validity at the individual level. We find that issue preferences measured via conjoint IMCEs strongly and consistently predict self-reported party support, vote intentions, and preferences on the same political issues. In addition to being reliably different from zero, the associations between conjoint-measured issue preferences and self-reported support for existing political parties are substantial in terms of size. Moreover, we show that IMCEs calculated from rating and choice outcomes predict relevant respondent-level covariates equally well.

When interpreting our findings, it is necessary to bear in mind that this paper is a purely empirical contribution. We demonstrate that one can recover individual-level estimates of preferences from conjoint experiments, and that these estimates reliably predict self-reported party support, vote intentions, and issue positions in expected ways. At the same time, our analyses do not speak to the broader conceptual debate on whether MCE-based quantities are the optimal preference estimands in conjoint experiments (Abramson et al. 2022; Bansak et al. 2022). Our results also do not imply that specific AMCE estimates always have signs or magnitudes that coincide with the median (majority) preferences.

Even given these caveats, our paper makes several important contributions to the methodological literature on conjoint experiments. The finding that conjoint-estimated preferences consistently predict self-reported attitudes and behavioral intentions has implications for the debate about their reliability. There is recent evidence that conjoint estimates incorporate large amounts of random noise—substantially larger than standard survey questions (Clayton et al. 2023). While our validity tests do not directly assess measurement error, they still suggest that conjoint MCEs are reliable enough to be useful measures of preferences. Since random error in predictor variables biases regression coefficients toward zero (King et al. 1994), our strong and consistent results are even more remarkable. The question to be addressed in future research is whether the measurement error is indeed random—but this is exactly what individual-level analyses of conjoint data, like those used in this paper, can provide an answer for.

Our results are relevant for the recent trend of using MCEs from conjoint experiments as unobtrusive measures of revealed preferences (Castanho Silva et al. 2023; Neuner and Wratil 2022). This usage is informed by the finding that conjoint experiments partially mitigate social desirability concerns (Horiuchi et al. 2022), somewhat similar to list experiments (Kuklinski et al. 1997). However, up to this point any mismatch between conjoint-estimated (revealed) and self-reported (expressed) preferences could have been attributed to potential issues with the measurement properties of MCEs. Our results confirm the predictive validity of MCEs as measures of preferences and thus suggest that whenever they diverge from self-reported preferences, social desirability may be at play.

The finding regarding equal predictive power of IMCEs based on rating and choice outcomes is important in light of recent work expressing methodological concerns about the latter. In addition to estimation issues like compositional effects in forced-choice designs (Ganter 2023), binary and rating outcomes may lead researchers to different substantive conclusions about preferences in the studied populations (Treger 2025). Our individual-level analyses, in contrast, demonstrate that conjoint-estimated preferences show predictive validity even when forced-choice outcomes are used.

In addition to demonstrating the predictive validity of conjoint MCEs using individual-level analyses, our results make a contribution regarding survey experiments more broadly. Their validity has been a prominent concern in the methodological literature (Barabas and Jerit 2010; Findley et al. 2017). Importantly, the fact that respondents in conjoint tasks are asked to choose from or rate hypothetical profiles can exacerbate the validity concern. However, we demonstrate that individual-level issue preferences estimated from conjoint experiments with hypothetical candidates predict self-reported support for existing parties. These results suggest that survey experiments—of which conjoint designs are a subclass—capture valuable information and thus are a useful method in the political science toolkit.

We also make a few narrower contributions to the practice of IMCE estimation. Feasible estimation of IMCEs requires minimizing the number of values per attribute, and this can be done at either the design or analysis stage. We use design-stage dichotomization of attribute values in the U.S. studies and analysis-stage dichotomization in the German study, and all studies return meaningful results at the individual level. This means that, in practice, minimization of values per attribute can be done at either stage. Moreover, we have been able to estimate IMCEs with choice outcomes and—in the German study—with only ten rated profiles per respondent. This runs counter to current recommendations for IMCE estimation (Zhirkov 2022), but our results suggest that those guidelines may be too restrictive and that IMCEs can be reliably estimated for a broader set of conjoint designs than previously thought.

Taken together, our results are encouraging for conjoint methodology in political science. Leveraging individual-level analyses, we demonstrate that MCEs—the main causal estimands in conjoint analyses—show predictive validity as measures of preferences. We would like to emphasize that our findings do not imply that researchers should not take seriously refinements to the interpretation and estimation of conjoint MCEs. Recent examples of such refinements include correction of measurement error (Clayton et al. 2023), adjustments for the real-world distributions of attributes (de la Cuesta et al. 2022), accounting for the presence of ties in forced-choice designs (Ganter 2023), and Bayesian estimation of IMCEs using machine learning (Robinson and Duch 2024). Rather, our results suggest that even unadjusted MCEs estimated via OLS regression, which are the quantities reported in most applied studies, have clear predictive validity and thus constitute useful measures of preferences.

References

Abramson, Scott F., Korhan Kocak, and Asya Magazinnik. 2022. "What Do We Learn about Voter Preferences from Conjoint Experiments?" *American Journal of Political Science* 66 (4): 1008–20.

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3): 529–46.

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

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

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2018. "The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments." *Political Analysis* 26 (1): 112–19.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2021. "Conjoint Survey Experiments." In *Advances in Experimental Political Science*, edited by James N. Druckman and Donald P. Green. Cambridge University Press.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2022. "Using Conjoint Experiments to Analyze Election Outcomes: The Essential Role of the Average Marginal Component Effect." *Political Analysis* 31 (4): 500–18.

Bansak, Kirk, and Libby Jenke. 2025. "Odd Profiles in Conjoint Experimental Designs: Effects on Survey-Taking Attention and Behavior." *Political Analysis*. Published ahead of print. https://doi.org/10.1177/2053168018822174.

Barabas, Jason, and Jennifer Jerit. 2010. "Are Survey Experiments Externally Valid?" *American Political Science Review* 104 (2): 226–42.

Carnes, Nicholas, and Noam Lupu. 2016. "Do Voters Dislike Working-Class Candidates? Voter Biases and the Descriptive Underrepresentation of the Working Class." *American Political Science Review* 110 (4): 832–44.

Castanho Silva, Bruno, Fabian Guy Neuner, and Christopher Wratil. 2023. "Populism and Candidate Support in the US: The Effects of 'Thin' and 'Host' Ideology." *Journal of Experimental Political Science* 10 (3): 438–47.

Chandler, Jesse, Cheskie Rosenzweig, Aaron J. Moss, Jonathan Robinson, and Leib Litman. 2019. "Online Panels in Social Science Research: Expanding Sampling Methods beyond Mechanical Turk." *Behavior Research Methods* 51 (5): 2022–38.

Clayton, Katherine, Yusaku Horiuchi, Aaron R. Kaufman, Gary King, and Mayya Komisarchik. 2023. "Correcting Measurement Error Bias in Conjoint Survey Experiments." Unpublished manuscript. https://gking.harvard.edu/conjointE.

Coppock, Alexander, and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research & Politics* 6 (1). https://doi.org/10.1177/2053168018822174.

Costa, Mia. 2021. "Ideology, Not Affect: What Americans Want from Political Representation." *American Journal of Political Science* 65 (2): 342–58.

de la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. 2022. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Political Analysis* 31 (1): 19–45.

Findley, Michael G., Brock Laney, Daniel L. Nielson, and J. C. Sharman. 2017. "External Validity in Parallel Global Field and Survey Experiments on Anonymous Incorporation." *Journal of Politics* 79 (3): 856–72.

Ganter, Flavien. 2023. "Identification of Preferences in Forced-Choice Conjoint Experiments: Reassessing the Quantity of Interest." *Political Analysis* 31 (1): 98–112.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments against Real-World Behavior." *Proceedings of the National Academy of Sciences* 112 (8): 2395–400.

Hainmueller, Jens, and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes toward Immigrants." *American Journal of Political Science* 59 (3): 529–48.

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

Hanretty, Chris, Benjamin E. Lauderdale, and Nick Vivyan. 2020. "A Choice-Based Measure of Issue Importance in the Electorate." *American Journal of Political Science* 64 (3): 519–35.

Horiuchi, Yusaku, Zachary Markovich, and Teppei Yamamoto. 2022. "Does Conjoint Analysis Mitigate Social Desirability Bias?" *Political Analysis* 30 (4): 535–49.

Imai, Kosuke, Bethany Park, and Kenneth F. Greene. 2015. "Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models." *Political Analysis* 23 (2): 180–96.

Jenke, Libby, Kirk Bansak, Jens Hainmueller, and Dominik Hangartner. 2021. "Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments." *Political Analysis* 29 (1): 75–101.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton University Press.

Kuklinski, James H., Michael D. Cobb, and Martin Gilens. 1997. "Racial Attitudes and the 'New South'." *Journal of Politics* 59 (2): 323–49.

Leeper, Thomas J., Sara B. Hobolt, and James Tilley. 2020. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis* 28 (2): 207–21.

Miller, David R., and Jeffrey Ziegler. 2024. "Preferential Abstention in Conjoint Experiments." *Research & Politics* 11 (4). https://doi.org/10.1177/20531680241299329.

Mummolo, Jonathan. 2016. "News from the Other Side: How Topic Relevance Limits the Prevalence of Partisan Selective Exposure." *Journal of Politics* 78 (3): 763–73.

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

Neuner, Fabian G., and Christopher Wratil. 2022. "The Populist Marketplace: Unpacking the Role of 'Thin' and 'Thick' Ideology." *Political Behavior* 44: 551–74.

Robinson, Thomas S., and Raymond M. Duch. 2024. "How to Detect Heterogeneity in Conjoint Experiments." *Journal of Politics* 86 (2): 412–27.

Treger, Clareta. 2025. "Changing the Lens: The Contingency of Results from Conjoint Experiments on the Outcome Variable and the Estimand." *Research & Politics* 12 (2). https://doi.org/10.1177/20531680251351238.

Zhirkov, Kirill. 2022. "Estimating and Using Individual Marginal Component Effects from Conjoint Experiments." *Political Analysis* 30 (2): 236–49.

Supplementary Information

Contents

Section A. Accounting for uncertainty in IMCE estimation

IMCEs are uncertain, because their estimation relies on relatively small numbers of observations. To account for uncertainty of the estimates, we use a procedure of drawing potential IMCE values multiple times from their estimated sampling distributions. This step relies on the assumptions that the OLS estimator used to estimate IMCEs is unbiased and normally distributed. Then, the regression with IMCEs as predictors is estimated independently for each such draw. Finally, results across estimated regressions are combined as if they were multiple imputations. The procedure is formalized below.

Consider the example where coefficient β_i is the point estimate of the IMCE for the value "Deport" of the candidate's position on immigration:

$$\text{rating}_{ij} = \alpha_i + \beta_i [\text{Immigrants} = \text{Deport}]_{ij} + \varepsilon_{ij}.$$

As soon as unbiasedness and normality assumptios hold, the sampling distribution of the corresponding OLS estimator is:

$$\hat{\beta}_i \sim \mathcal{N}\left(\beta_i,\ \text{Var}(\hat{\beta}_i)\right).$$

We draw 100 potential values of β_i, and thus the corresponding IMCE, from the estimated sampling distribution with the mean equal to the point estimate and the variance equal to squared standard error:

$$\hat{\beta}_{i,1}, \ldots, \hat{\beta}_{i,100} \sim \mathcal{N}\left(\hat{\beta}_i,\ \text{SE}(\hat{\beta}_i)^2\right).$$

Each $\hat{\beta}_{i,m}$ with $m \in \{1, \ldots, 100\}$ is a plausible value of the IMCE for the value "Deport" of the candidate's position on immigration. They can be used in separate 100 regression models to estimate the effect of this issue position on candidate support:

$$\text{support}_i = \gamma_m + \delta_m (\text{IMCE: Immigrants, deport})_{i,m} + u_{i,m}.$$

Then, the overall coefficient point estimate is the average across the regressions:

$$\hat{\delta} = \frac{1}{100} \sum_{m=1}^{100} \hat{\delta}_m.$$

And its total variance is the combination of within-estimation and between-estimation variances according to Rubin's rules:

$$\widehat{\text{Var}}(\hat{\delta}) = \left[\frac{1}{100} \sum_{m=1}^{100} \text{SE}(\hat{\delta}_m)^2\right] + \left(1 + \frac{1}{100}\right) \left[\frac{1}{99} \sum_{m=1}^{100} (\hat{\delta} - \hat{\delta}_m)^2\right].$$

Section B. Sample candidate profiles in the conjoint task, Studies 1 and 2

Please carefully review the information about the candidates presented below, then answer the questions.

	Candidate 1	**Candidate 2**
Race	Asian	Hispanic
Gender	Female	Male
Age	63	64
Position on unauthorized immigrants	Send unauthorized immigrants back to their home countries	Allow unauthorized immigrants to remain in the United States
Position on the size of government	Government should provide more services in areas like health and education, even if it means raising taxes	Government should provide fewer services in areas like health and education to reduce spending and cut taxes
Position on abortion	Abortion should not be permitted	A woman should be able to obtain an abortion

Which of the two candidates would you rather vote for?

Candidate 1

Candidate 2

On a scale from 0 to 10 where 0 means that you definitely would not consider voting for the candidate and 10 means that you definitely would consider voting for the candidate, how would you rate Candidate 1 and Candidate 2?

Definitely would not consider voting for — Definitely would consider voting for

0 1 2 3 4 5 6 7 8 9 10

Candidate 1

Candidate 2

Section C. Survey questions, Studies 1 and 2

Feeling thermometers

We would like to get your feelings toward the two main political parties using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the party. Ratings between 0 degrees and 50 degrees mean that you do not feel favorable toward the party and that you do not care too much for that party. You would rate the party at the 50 degree mark if you do not feel particularly warm or cold toward the party.

- Democratic Party
- Republican Party

Presidential vote intention

Thinking about the 2024 presidential election, are you more likely to vote for:

- Joe Biden, the presumptive Democratic nominee
- Donald Trump, the presumptive Republican nominee

Generic congressional ballot

Thinking about voting in your congressional district, are you more likely to vote for:

- A Democratic candidate
- A Republican candidate

Position on the size of government

Some people think the government should provide fewer services even in areas such as health and education in order to reduce spending and cut taxes. Other people feel it is important for the government to provide many more services even if it means an increase in taxes. What is your opinion?

- Government should provide fewer services
- Government should provide more services

Position on abortion

There has been some discussion about abortion in recent years. Which of the following opinions comes closest to your view?

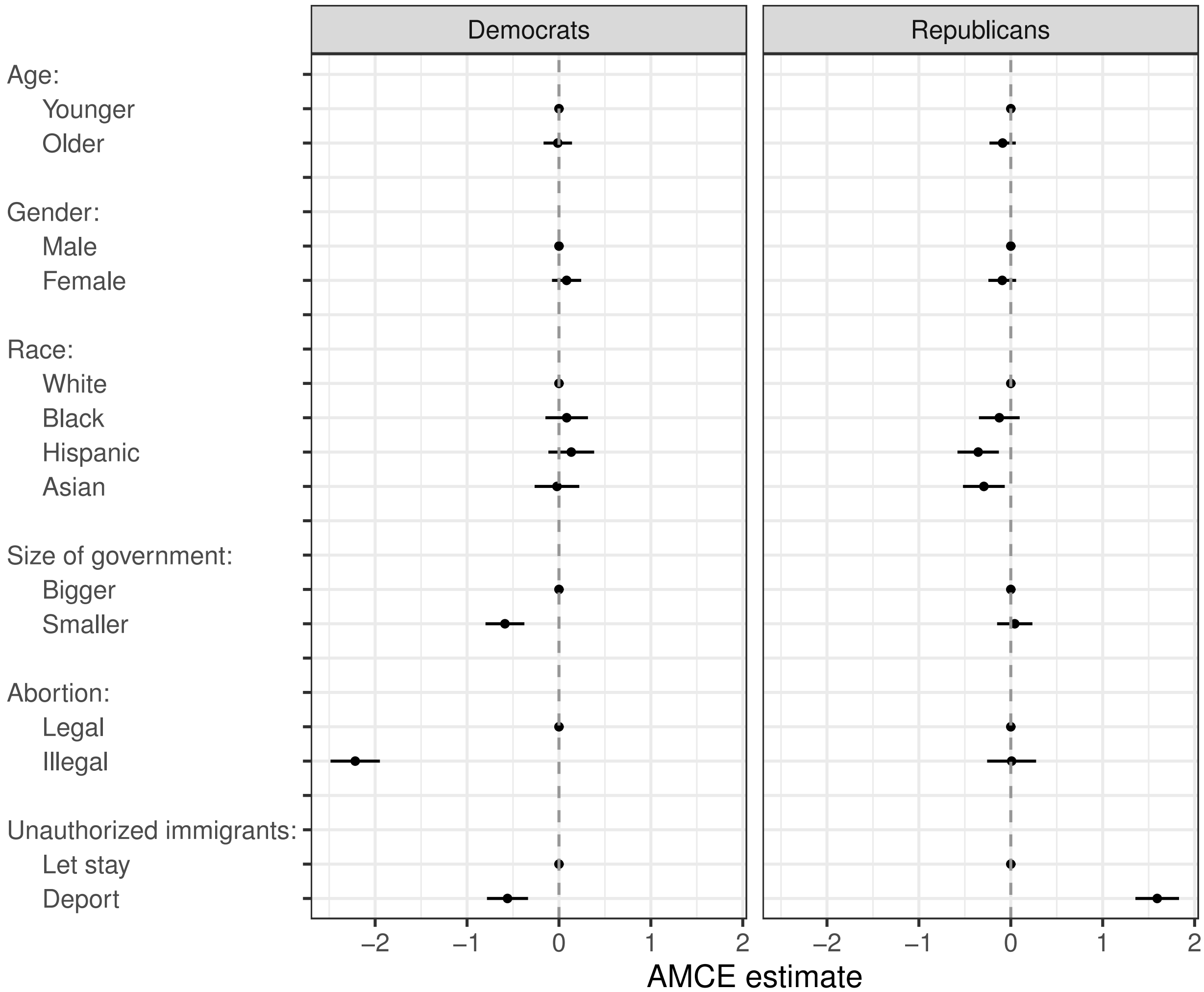
- In most cases, abortion should not be permitted
- In most cases, a woman should be able to obtain an abortion

Position on unauthorized immigrants

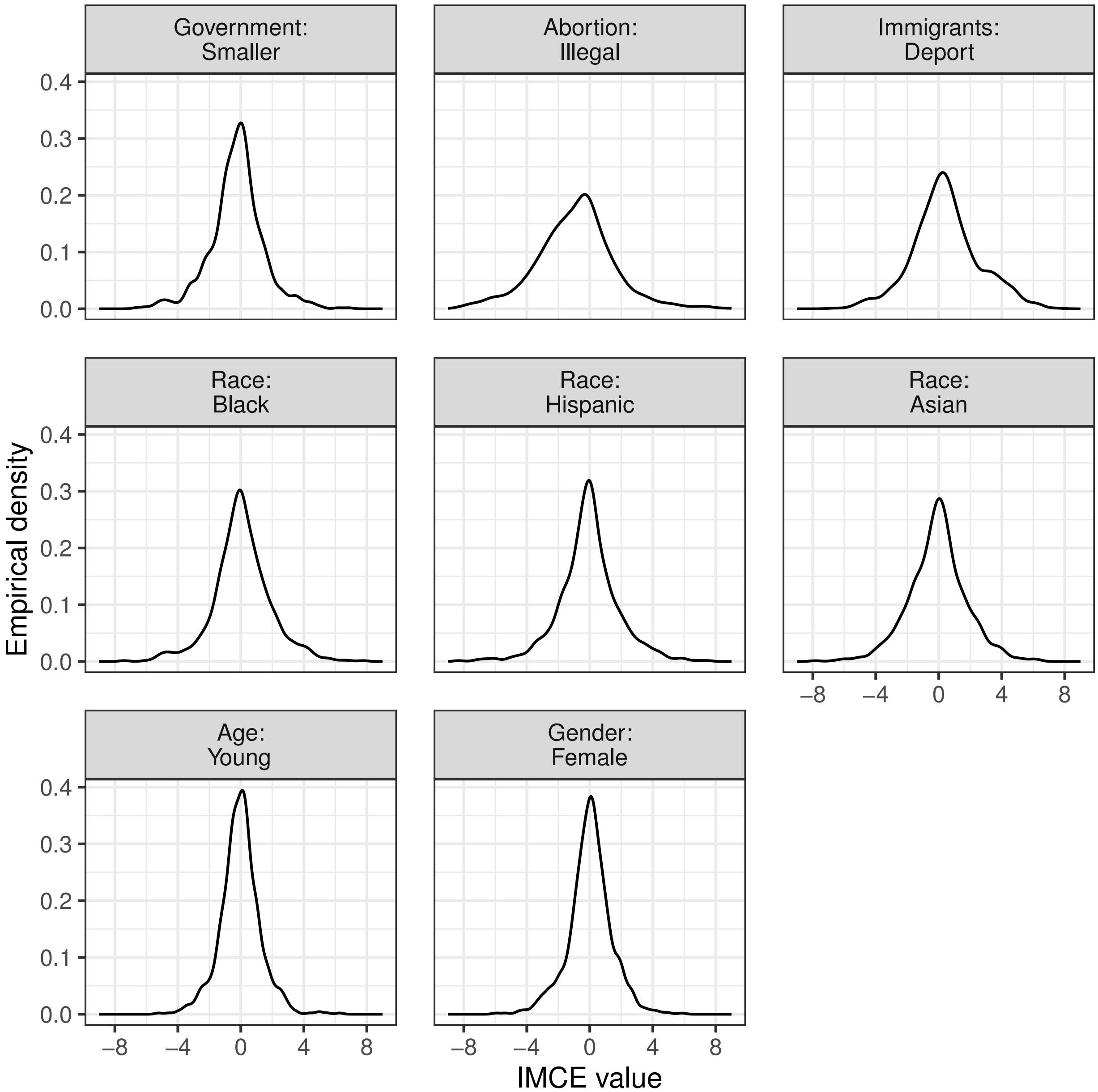
Which comes closest to your view about what government policy should be toward unauthorized immigrants now living in the United States?

- Send unauthorized immigrants back to their home countries
- Allow unauthorized immigrants to remain in the United States

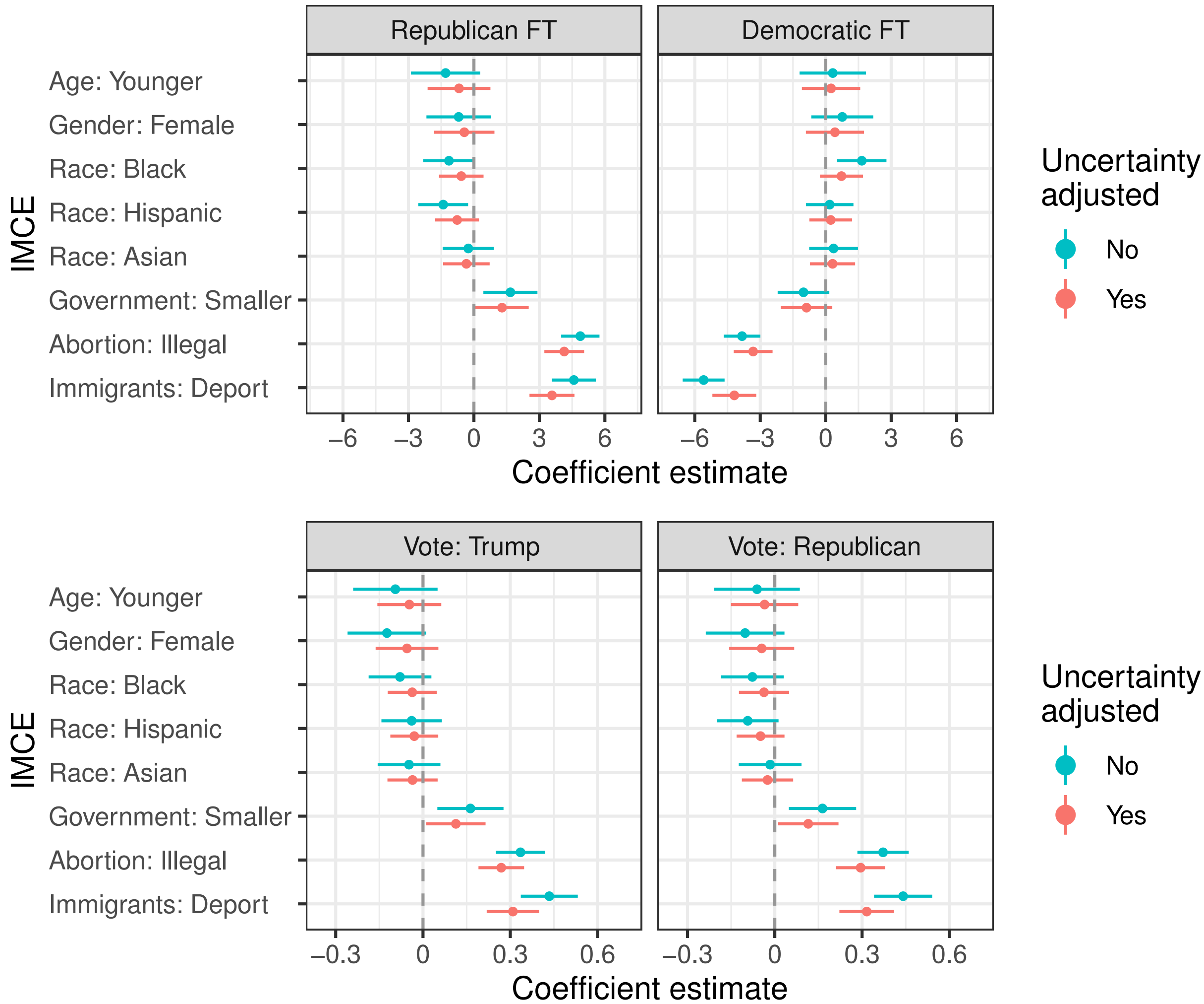
Section D. AMCEs by respondent's party, Study 1



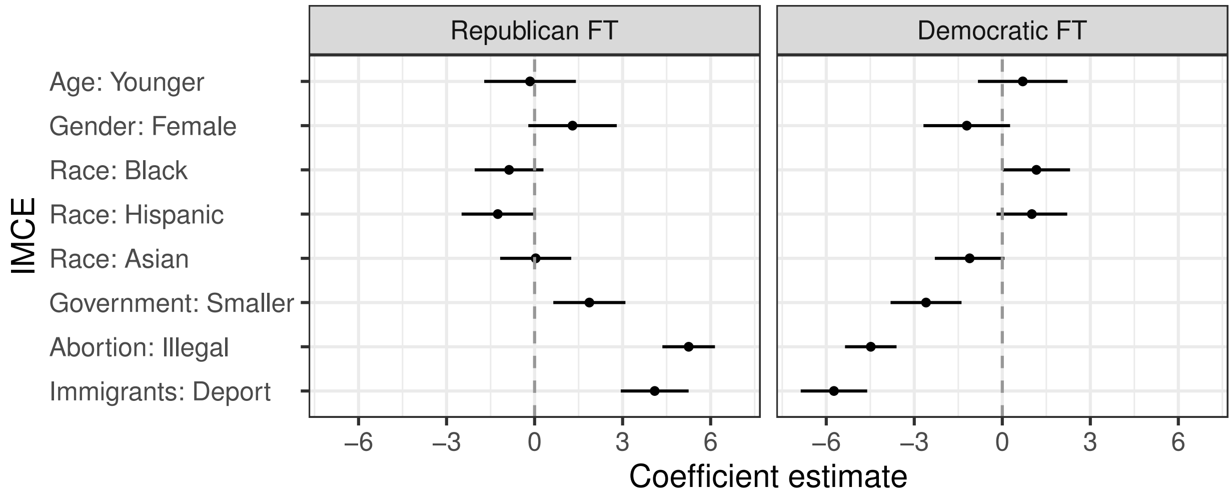
Section E. Empirical densities of estimated IMCEs, Study 1

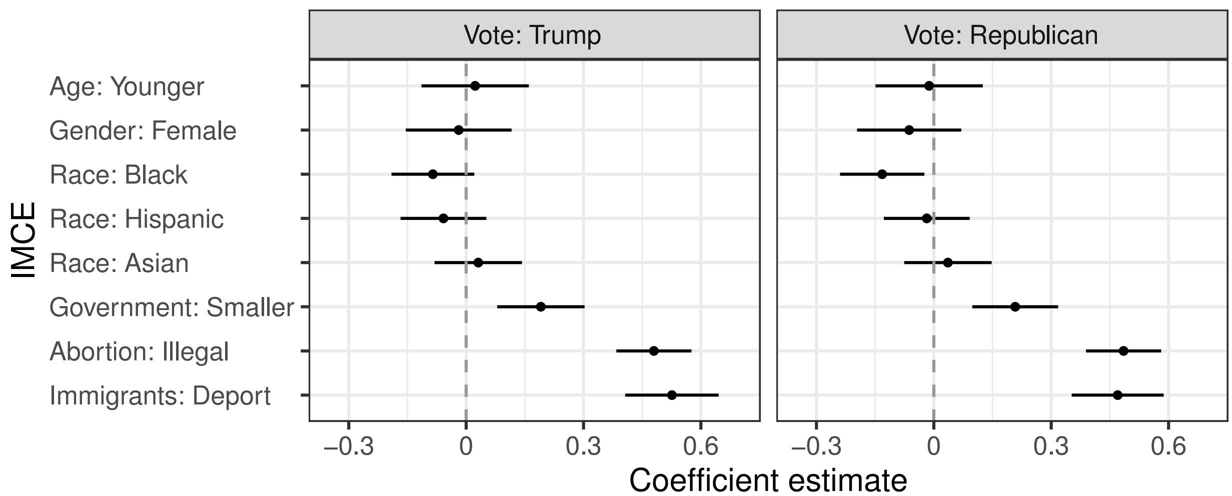


Section F. Uncertainty-adjusted estimates, Study 1

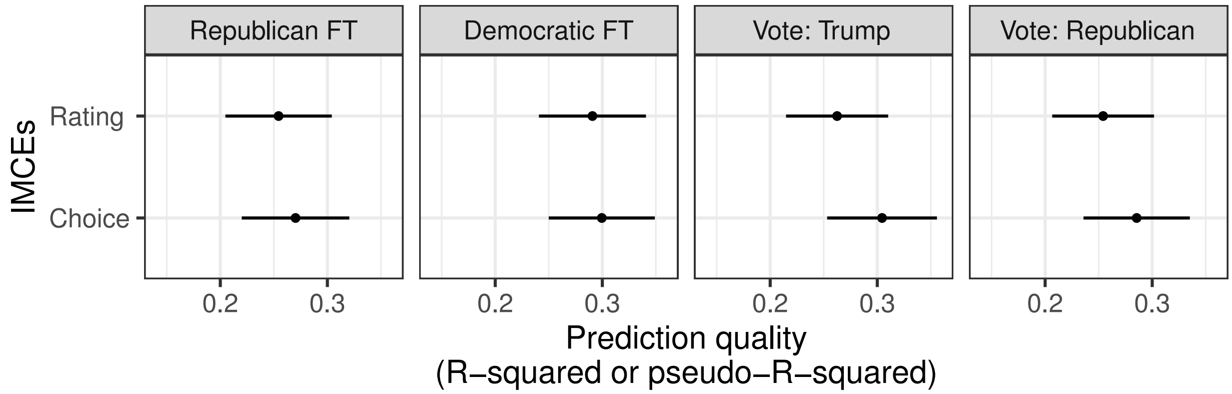


Section G. Predictive validity, Study 2

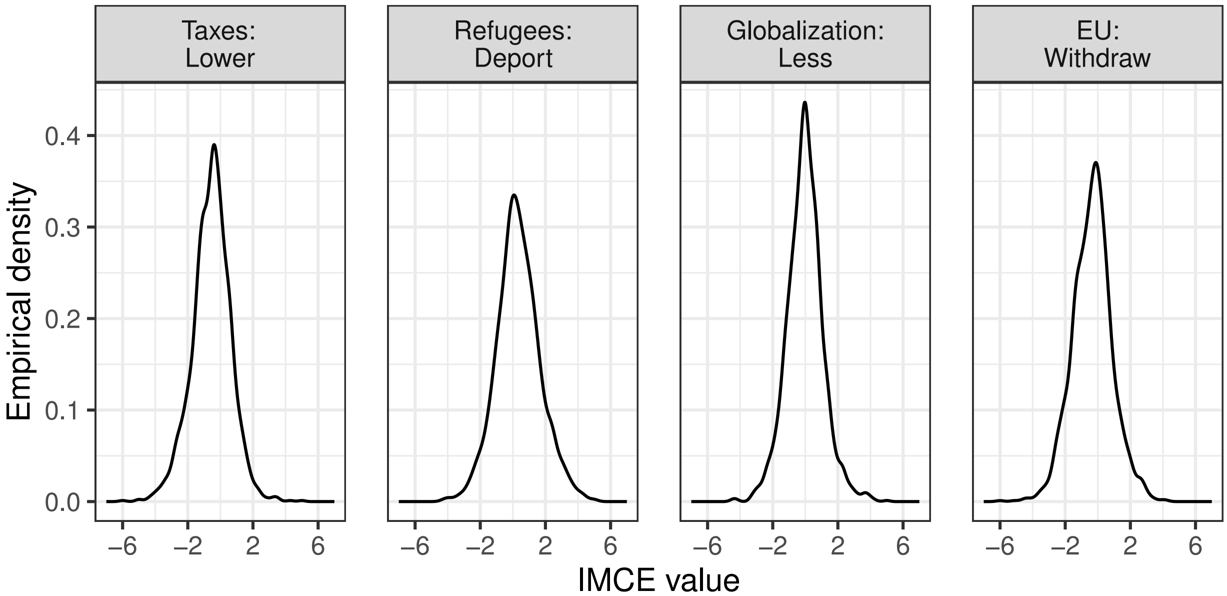




Section H. Performance of rating and choice outcomes, Study 2



Section I. Empirical densities of estimated IMCEs, Study 3



Section J. Uncertainty-adjusted estimates, Study 3