

Supplemental materials for: What Do We Learn From a Decade of Candidate Choice Conjoint Experiments?

Alexander Coppock, Matthew Blyth, Marco M. Aviña, and Semra Sevi*

October 19, 2026

Contents

A	Software and interactive tool	3
B	Search and screening	7
C	Study manifest	24
D	Generalizability	33
E	Validation against published estimates	37
F	Additional regression tables	39
	F.1 Heterogeneity by respondent partisanship	39
	F.2 Heterogeneity by presence of policy information	41
G	Attenuation in the number of attributes	51
H	Recoverable no-choice rates	55
I	Satisficing correction subset robustness	56

*Alexander Coppock is Associate Professor of Political Science at Northwestern University. Matthew Blyth is a Tobin Center pre-doctoral fellow at Yale University. Marco M. Aviña is a doctoral candidate in Government and Social Policy at Harvard University. Semra Sevi is Assistant Professor of Political Science at the University of Toronto.

J	AFCP intransitivity audit	58
K	Publication bias	60
L	Party–policy tradeoffs by respondent ideology	62
M	Party–policy tradeoffs by respondents’ own issue positions	64
N	Party-demographic forest plots	66
O	Per-attribute forest plots	79
P	Covariate-adjusted AFCPs	104
Q	Estimates underlying the main-text figures	107
R	Use of AI tools	122

A Software and interactive tool

The main text’s Discussion promises two software artifacts to accompany the paper; this section documents both: the matchup-extraction R package and the interactive AFCP Explorer.

Extracting order-invariant matchups

Carrying out this reanalysis at scale required one operation that existing conjoint software does not provide. For any pair of candidate attribute levels, we need every task in which the two profiles differ on exactly those levels, together with a record of which candidate was chosen, regardless of the order in which the two profiles happened to be presented. Standard tools treat the profile position (Candidate 1 versus Candidate 2, left versus right) as fixed and estimate marginal component effects from it; they do not return an order-invariant, matchup-level binary outcome that can be pooled across studies whose attribute schemes differ. We therefore wrote a function, `get_matchups()`, that takes a conjoint dataset and two attribute-level specifications (a Candidate A and a Candidate B, each of which may constrain several attributes at once) and returns a tidy table of the tasks in which one profile matches A and the other matches B, in either presentation order, along with a single indicator for whether the A-side candidate was chosen. Because the function is agnostic to candidate order and to the number of constraining attributes, the same call produces the single-attribute matchups (for example, a woman candidate versus a man candidate) and the multi-attribute matchups used throughout the paper. We make this software freely available as an open-source R package, `conjointmatchups` (<https://github.com/acoppock/conjointmatchups>), with documentation and worked examples.

The AFCP Explorer

We also provide an interactive companion, the AFCP Explorer, a web application that lets a reader assemble any two candidate profiles and read the meta-analytic Average Feature Choice Probability that Candidate A is chosen over Candidate B, pooled across every experiment that randomized the relevant attributes and, optionally, split by respondent characteristics. Each estimate is recomputed on demand from the standardized task-level data using the same pipeline as the paper, so the paper’s pooled estimates can be reproduced closely and inspected directly. Figure A.1 shows the application reproducing the headline gender result, and Figure A.2 shows the Republican-versus-Democrat matchup split by respondent

partisanship. The application is available at <https://anonymouspoliticalscientist.shinyapps.io/afcp-explorer/>; the public version deposited at publication will include whatever underlying data the original study authors permit us to redistribute.

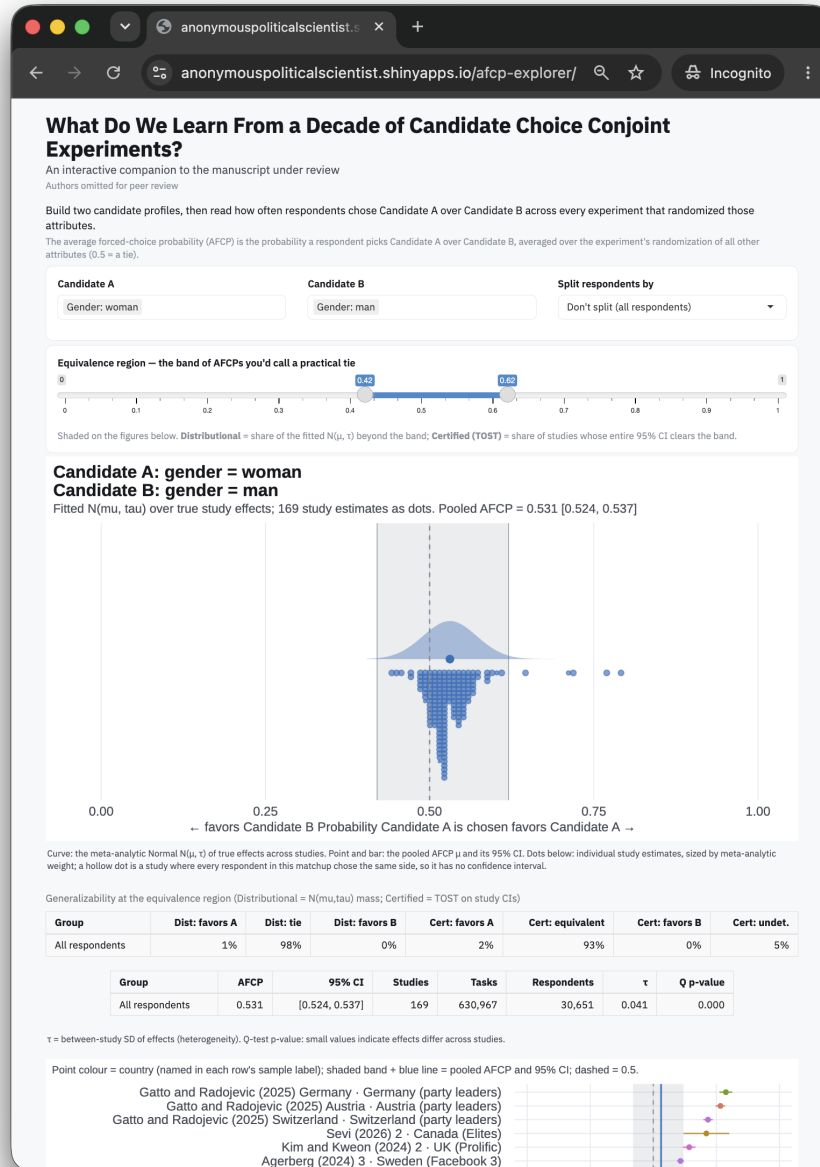


Figure A.1: The AFCEP Explorer reproducing the pooled woman-versus-man result
Note: Estimates cover conjoints from all countries, not US studies only. Output of the AFCEP Explorer for Candidate A = Woman versus Candidate B = Man. The pooled Average Feature Choice Probability is 0.530 [0.525, 0.536] across the 212 studies that randomize candidate gender. The curve is the fitted $N(\mu, \tau)$ over true study effects; each dot is one study, sized by its meta-analytic weight; the dashed line marks the 0.5 point of indifference.



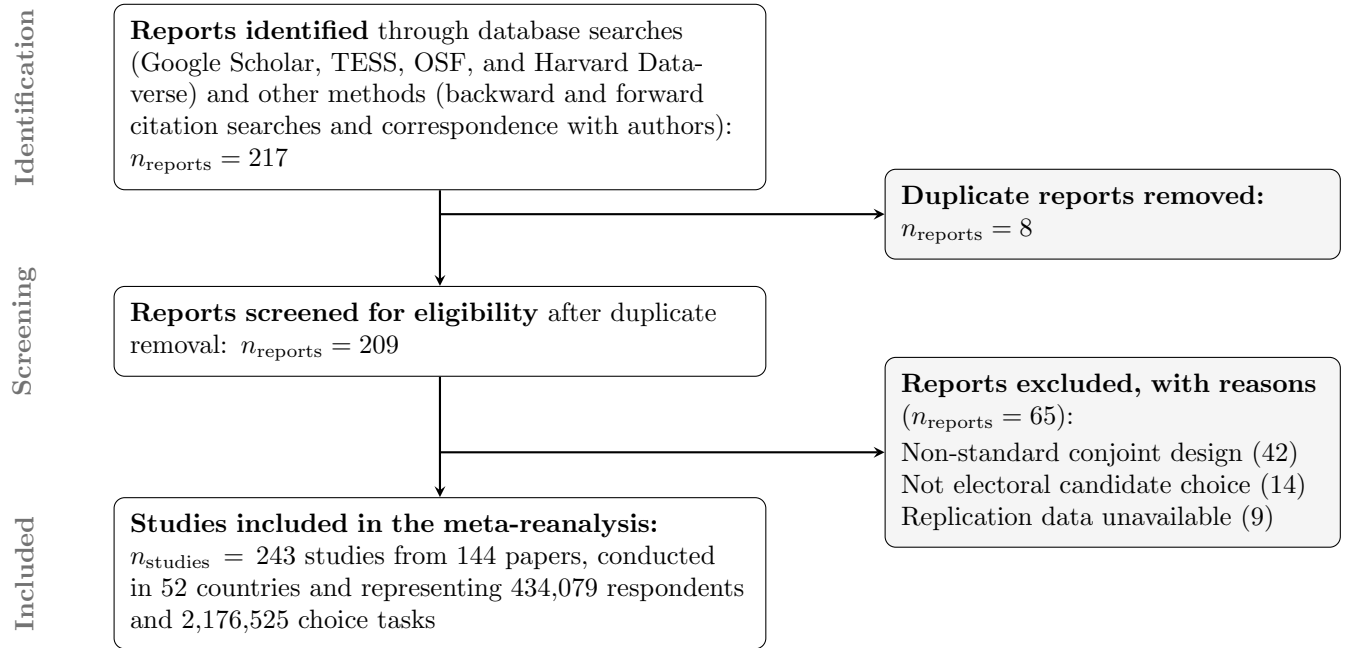
Figure A.2: The AFCEP Explorer reproducing the Republican-versus-Democrat matchup by respondent partisanship

Note: Output of the AFCEP Explorer for Candidate A = Republican versus Candidate B = Democrat, split by the respondent's own party. Republican respondents choose the Republican candidate 72.1% of the time and Democratic respondents 27.4% of the time, while Independents are near indifference. Curves are the fitted $N(\mu, \tau)$ over true study effects within each respondent group; dots are studies sized by meta-analytic weight; the dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted. The Explorer pools one estimate per study, while the main text estimates separately the policy-shown and policy-hidden arms of the four studies that randomize whether policy positions appear, so its values (72.4% and 27.2%) differ slightly from these.

B Search and screening

The corpus was assembled through the search, screening, and inclusion process summarized in Figure B.3. The main text’s “Meta-reanalysis design” section describes this process in brief and points here for the PRISMA diagram.

Figure B.3: PRISMA 2020 flow diagram for study identification, screening, and inclusion



Note: Included studies are the cleaned corpus analyzed in this paper; excluded reports are the reviewed-and-rejected studies from our screening records, consolidated into three reason categories. Counts in the identification, screening, and exclusion boxes are reports (papers); the inclusion box reports both studies and papers. A paper contributing at least one study to the corpus counts as an included report, so the handful of papers with one qualifying study and one rejected study appear only on the included side; their rejected studies are listed in the appendix table of excluded studies. Because the corpus was assembled through iterative searching across databases and registers (Google Scholar, TESS, OSF, and Harvard Dataverse), backward and forward citation searches, and correspondence with authors, the identification count reflects reports brought forward for eligibility assessment rather than an enumeration of all database hits. Replication data were requested directly from authors for 36 reports.

Table B.1 lists every study reviewed and rejected, with the reason for exclusion. The reasons are consolidated with the same mapping the PRISMA diagram uses, so the two cannot disagree.

Table B.1: Studies Reviewed but Excluded

Study	Reason	Detail
Funck & McCabe (2022) study 3	Duplicate report	Same conjoint experiment as <code>funck_mccabe_2022_study_1</code> . The two cleaning scripts read two copies of one dataset — <code>candlongAP.csv</code> and <code>candlongfin.csv</code> hold identical 12,810-row contents, the latter differing only by two fewer columns — so the cleaned outputs agree on the outcome and all 20 shared attributes. The paper reports a single conjoint experiment; the randomization of whether policy information is shown is a task-level variable within it (<code>task_policy_shown</code>), not a separate study.
Teele, Kalla & Rosenbluth (2018), Study 3	Duplicate report	The same experiment as <code>teele_kalla_rosenbluth_2018_study_1</code> : all 6,225 of <code>study_3</code> 's task fingerprints match <code>study_1</code> 's. <code>study_1</code> is the raw TESS3 GfK file and strictly dominates, with 2,144 respondents and 6,432 tasks against 2,105 and 6,225, plus two attributes <code>study_3</code> lacks (<code>cand_resp_age_3_match</code> , <code>cand_number</code>); <code>study_3</code> is the authors' extract with 39 respondents excluded. The <code>conjoint_studies.xlsx</code> row also mislabeled it as the legislators replication, which it is not. The paper stays in the corpus through studies 1 and 2. Dropped 2026-08-19; cleaned files in <code>clean_data/deduped_orphans/</code> .
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Abrajano, Elmendorf & Quinn (2018), Study 1	Non-standard joint design	Only candidate ethnicity is randomized; the other four attributes are fixed within six preselected candidate pairs, so the study fails the at-least-two-randomized-attributes rule. Found by the 2026-09-28 reproduction sweep (code/reproductions/abrajano_elmendorf_quinn_2018/NC) excluded 2026-09-28. The paper leaves the corpus.
Adiguzel (2025)	Non-standard joint design	Governance 2025; Turkey. Only two attributes, both policy positions (judiciary law, social policy); no candidate attribute the meta-analysis uses (holliday_rudkin_2025 precedent). Found in the IRW screen (2026-10-17).
Agerberg, M. (2024)	Non-standard joint design	EPSR 2024; 25 EU countries; mayoral candidate conjoint includes abstention option (A / B / neither) — not strictly binary forced-choice.
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason		Detail
Albert & Costa (2025) study 2	Non-standard joint design	con-	Only one profile is randomized. Candidate B is projected to lose the general election in all 15,924 tasks while candidate A is 'likely to win' or 'toss-up', and B's policy positions are not randomized either: they mirror the respondent's own stated positions. That is a design holding one profile's attributes constant, which the standard-design criterion excludes. Study 1 of the same paper randomizes both profiles and is retained. Separately, the cleaning script was found to misread the raw columns: support1/support2 are candidate A's stances on the two issues rather than the two candidates' stances, and issuejjanswer is the respondent's own position. A corrected script is preserved in cleaning_code/cleaning_archive/ should this inclusion decision ever be revisited.
Busby (2022)	Non-standard joint design	con-	Open-ended attribute evaluations; not binary candidate choice
Christensen & Saikkonen (2025, JEPOP)	Non-standard joint design	con-	Respondents evaluate tolerance of democratic norm violations by policy issue (Does Democratic Backsliding Depend on the Policy Issue?), not choosing among candidates for electoral office. Ineligible under standard design criteria.
Cozza et al. (2025)	Non-standard joint design	con-	BJPS 2025; UK; party leader evaluation conjoint; outcome is rating (deservingness/competency), not binary forced-choice vote choice.
Crowder-Meyer et al. (2020)	Non-standard joint design	con-	Replication data lacks profile and task identifiers; cannot reconstruct which two candidate profiles were paired in each choice task.
<i>Continued on next page</i>			

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason		Detail
Eggers (2018)	Non-standard joint design	con-	Same design issue as Eggers et al. (2017): paired candidates drawn from structurally different pools (incumbent vs. challenger).
Eggers et al. (2017)	Non-standard joint design	con-	Candidate profiles drawn from different distributions: one is always an incumbent and the other a challenger, violating the assumption of comparable candidates.
Frederiksen (2022)	Non-standard joint design	con-	Outcome is likelihood of support on a 1-5 Likert scale, not a binary forced-choice between two candidates.
Frederiksen (2024, JOP)	Non-standard joint design	con-	Journal of Politics 86(2); conjoint in the US, UK, Czech Republic, Mexico and South Korea. Respondents rate each candidate separately: the outcome ‘support’ is a 1-5 likelihood-of-support Likert item, not a binary forced choice between the two profiles. Same design family as the already-excluded frederiksen_2022 (APSR), which is a different paper. Replication files held at original_materials/1_archive/frederiksen_2024/.
Frederiksen & Laustsen (2025, BJPS)	Non-standard joint design	con-	British Journal of Political Science 55, e138; two US conjoint experiments on interstate conflict and undemocratic candidates. Respondents rate each candidate separately: the outcome ‘vote’ is a 7-point rating rescaled to 0-1, not a binary forced choice. Replication files held at original_materials/1_archive/frederiksen_laustsen_2025/.
Green, Schaffner & Luks (2022)	Non-standard joint design	con-	Only candidate policy position randomized
Gutierrez-Romero & Lebas (2020)	Non-standard joint design	con-	Only randomized candidate attribute is campaign violence history; remaining candidate characteristics are fixed across profiles.

Continued on next page

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason		Detail
Hall, Leigh & Solomon (2026, Nat Hum Behav)	Non-standard joint design	con-	Nature Human Behaviour 10: 896-905; US conjoint on democratic neutrality. Respondents rate each candidate separately: the outcome ‘votepref’ is a 1-5 rating (the authors’ ‘relvotepref’ is the difference between the two ratings), not a binary forced choice. Replication files held at original materials/1_archive/hall.leigh.solomon.2026/.
Hanretty, Lauderdale & Vivyan (2020)	Non-standard joint design	con-	Only policy positions randomized; 3 of 34 issues shown per task with numeric codes; attribute-issue mapping undecodeable from replication data
Harden & Layman (2022)	Non-standard joint design	con-	CCES module; unclear paired forced-choice design
Hartmann (2026), Study 1	Non-standard joint design	con-	All seven randomized attributes describe the incumbent’s disaster-policy conduct; none is a candidate characteristic. Same reason as holliday_rudkin.2025. IRW table screened at wave 1 ingest; excluded by Alex 2026-10-17.
Henderson et al. (2022)	Non-standard joint design	con-	Each task presents 8 candidate profiles from which respondents choose 4; not a standard binary forced-choice design.
Himmelroos, Christensen & Vento (2024)	Non-standard joint design	con-	Data lacks profile and task number identifiers; cannot reconstruct which candidates were paired in each choice task.
<i>Continued on next page</i>			

Table B.1: Studies Reviewed but Excluded (continued)

Study			Reason	Detail	
Hobolt	and	Os-	Non-standard	con-	Political Behavior 2025; UK; candidate conjoint about authoritarian behavior. The controversy type is a candidate attribute, but the remaining 4 attributes (reaction type, actor party, actor status, actor gender) all describe other MPs reactions to the candidate — not the candidate own characteristics. Does not meet candidate-attribute conjoint criterion.
nabrugge	(2025)		joint design		
Holliday	&	Rudkin	Non-standard	con-	Policy bundle conjoint ("DC On My Mind", 2025). Meets eligibility criteria on paper (2 profiles, binary forced choice, electoral candidates, ≥ 2 attributes) but has no candidate demographic attributes (no gender, race, age, experience). The only meta-analysis for which it would qualify is the respondent-candidate policy match analysis, and even that is a poor fit: deriving cand_policy_overall requires manually coding 25 policy positions as liberal/conservative (several are ideologically ambiguous), and each candidate receives only 4 policies per task, so most profiles land as "Mixed" with little variation in the Solid liberal/conservative cells.
(2025)			joint design		
<i>Continued on next page</i>					

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Håkansson & Grahn (2026), Study 1	Non-standard joint design	con- The ingested data (steg == 1 of the replication file, 8,382 party members with member and ideology attributes) are not the paper’s conjoint, and no paper reporting them could be identified. The paper’s conjoint (steg == 2, 1,276 party selectors x 5 tasks) is the experiment already in the corpus as grahn_thisell_2025_study_1. Found by the 2026-09-22 manifest audit; excluded 2026-09-23. The paper leaves the corpus. Cleaning script in <code>cleaning_code/cleaning_archive/</code> , cleaned files in <code>clean_data/excluded_studies/</code> .
Igarashi & Ono (2025)	Non-standard joint design	con- Single-profile vignette experiment (one candidate shown per respondent); no forced-choice comparison.
Igarashi and Ono (2025)	Non-standard joint design	con- JEPS 2025; Japan; vignette experiment measuring voting intentions via rating scale, not binary forced-choice conjoint.
Jankowski (2023)	Non-standard joint design	con- Survey experiment about a single named candidate (Maassen, 2021 German election); no conjoint attributes randomized across profile pairs.
Jankowski & Rehmert (2025)	Non-standard joint design	con- List deselection task (respondents exclude candidates from a set of 8); not a paired binary choice.
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study			Reason	Detail
Karpowitz	et al.	(2024), Study 1	Non-standard joint design	con- Candidate sex is never randomized (every task is Julie Baker versus Steven Hall); the randomized content is a self-presentation statement and occupation, so the study fails the at-least-two-randomized-candidate-attributes rule. Paper Study 3 (2016 GOP caucus attenders). IRW wave 1 ingest; excluded by Alex 2026-10-17.
Karpowitz	et al.	(2024), Study 2	Non-standard joint design	con- Candidate sex is never randomized (every task is Julie Baker versus Steven Hall); the randomized content is a self-presentation statement and occupation, so the study fails the at-least-two-randomized-candidate-attributes rule. Paper Study 4 (2018 GOP caucus attenders). IRW wave 1 ingest; excluded by Alex 2026-10-17.
Lust	& Benstead	(2024)	Non-standard joint design	con- Single candidate evaluated on a 0-10 favorability scale; no paired comparison or binary forced choice.
Lust	and Benstead	(2024)	Non-standard joint design	con- CPS 2024; six Arab countries; rating of likelihood to vote for candidate, not binary forced-choice between two profiles.
Martin	& Blinder	(2021)	Non-standard joint design	con- Replication data lacks respondent ID and task ID; cannot identify individual respondents or reconstruct paired choice tasks.
McLean & Ono (2024)			Non-standard joint design	con- exp1 has a single attribute (candidate age: Younger/Middle-Aged/Older) — too minimal for meta-analysis. exp2 win variable does not sum to 1 per respondent pair (1605 respondents have win=0 for both candidates), so it is not a standard forced-choice binary outcome.
<i>Continued on next page</i>				

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason		Detail
Miwa (2025)	Non-standard joint design	con-	Rating scale outcome; candidates rated individually rather than selected in a forced-choice comparison.
Miwa, Kasuya & Ono (2023)	Non-standard joint design	con-	Rating scale outcome (1-8); candidates are rated rather than chosen in a forced-choice comparison.
Nadler, Hepplewhite & van Oosten (2025)	Non-standard joint design	con-	Sequential individual evaluation of 8 politicians followed by retrospective pairwise choice (EMMAVID design); not a standard forced-choice binary outcome.
Pedersen & Dahlgaard (2019)	Non-standard joint design	con-	Rating scale outcome (1-6 per candidate per attribute); not a binary forced-choice between two complete profiles.
Reher & Evans (2025)	Non-standard joint design	con-	4 candidates presented per task; not a standard paired (2-candidate) forced-choice design.
Schneider (2020)	Non-standard joint design	con-	Data lacks profile and task number identifiers; cannot reconstruct which candidates were paired in each choice task.
Titelman and Lauderdale (2024)	Non-standard joint design	con-	PSRM 2024; UK candidate conjoint; candidate attributes not independently randomized — profiles drawn from real Conservative/Labour candidate distributions, preserving real-world covariance. Primary treatment is party label presence/absence, not a standard AMCE conjoint.

Continued on next page

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Torcal (2026), EPADE Chile	Non-standard joint design	con- Data-only release with no write-up: the EPADE panel (Torcal, Universitat Pompeu Fabra; OSF doi:10.17605/osf.io/u43pb, project created 2026-03-04), candidate-preference conjoint for head of government in Chile. Excluded under the three-year rule for data-only studies (fieldwork ended less than three years before the 2026-10-16 search); revisit at the next search. IRW table torcal_2026_epade_candidates_chile. Ruled by Alex 2026-10-17.
Torcal (2026), EPADE Germany, Portugal, Netherlands	Non-standard joint design	con- Data-only release with no write-up: the EPADE panel (Torcal, Universitat Pompeu Fabra; OSF doi:10.17605/osf.io/u43pb, project created 2026-03-04), candidate-preference conjoint for head of government in Germany, Portugal, and the Netherlands. Excluded under the three-year rule for data-only studies (fieldwork ended less than three years before the 2026-10-16 search); revisit at the next search. IRW table torcal_2026_epade_candidates. Ruled by Alex 2026-10-17.
Trexler & Johnston (2025)	Non-standard joint design	con- Political Behavior 2025; United States. Outcome is a 4-point preference scale, not a binary forced choice (tomz_2019 precedent). Found in the IRW screen (2026-10-17).
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Yang (2026), Study 1	Non-standard joint design	con- Candidate gender is never randomized within a pair (every task is one woman against one man) and post type (Campaign, Personal Viewpoints, Personal Lives) is assigned by five between-respondent groups, so the study does not identify a gender effect; the control baseline also carries the gender-coloured posts the paper flags. Found by the 2026-10-18 reproduction sweep (code/reproductions/archive/yang_2026/NOTES.md); excluded 2026-10-18. The paper leaves the corpus. Cleaning script in cleaning_code/cleaning_archive/, cleaned files in clean_data/excluded_studies/.
van Oosten (2023)	Non-standard joint design	con- Same EMMAVID sequential evaluation + retrospective pairwise choice design as Nadler, Hepplewhite & van Oosten (2025).
van Oosten (2025)	Non-standard joint design	con- Same EMMAVID sequential evaluation + retrospective pairwise choice design as Nadler, Hepplewhite & van Oosten (2025).
Agadjanian et al. (2023), Study 1	Not electoral candidate choice	Hiring conjoint for an appointed city manager; the paper’s other arm hires a biology professor. Mechanically eligible (exactly 2 profiles per task, 1 selected, 2 or more randomized attributes) but the profiles are applicants for an appointed post, so it fails the criterion that respondents choose among candidates for electoral office. Cleaned and in the corpus until 2026-08-19; cleaning script in cleaning_code/cleaning_archive/, cleaned files in clean_data/excluded_studies/.
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Arnesen, Duell & Johannesson (2019), Study 2	Not electoral candidate choice	The outcome asks respondents to guess which of the two candidates holds a stated policy position; it is not a choice between candidates. Mechanically eligible, so no check caught it; found by the 2026-09-22 manifest audit. The paper stays in the corpus through study 1. Cleaning script in <code>cleaning_code/cleaning_archive/</code> , cleaned files in <code>clean_data/excluded_studies/</code> .
Blumenau, Wolkenstein & Wratil (2025)	Not electoral candidate choice	Perspectives on Politics 2025; Germany, United Kingdom, and United States. The forced choice asks which politician would better represent the respondent, not which they would vote for. Found in the IRW screen (2026-10-17).
Brehm, Zhou & Denney (2025)	Not electoral candidate choice	Communist and Post-Communist Studies 2025; Germany, South Korea, and North Korean migrants in South Korea (three samples). Respondents read the opinions of two hypothetical "emerging politicians" on democracy, markets, welfare, gender, and diversity and answer "With whom do you most agree?" (Table 3). No office or election; agreement, not vote choice; no candidate-characteristic attributes. Found in the IRW wave-2 ingest (2026-10-18).
Devine, Turnbull-Dugarte & Ryan (2025)	Not electoral candidate choice	Outcome measures class perceptions and electoral success predictions
Evert and Klint (2019)	Not electoral candidate choice	Unpublished student thesis (University of Copenhagen); outcome is political trust rating, not binary candidate choice.
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Fernandes et al. (2025)	Not electoral candidate choice	BJPS 2025; France/Israel/Sweden; choice task is between two fictional parties (Party A/B), not two candidates. All five attributes are party-level. Ineligible on candidate choice criterion.
Higashijima et al. (2025)	Not electoral candidate choice	Political Behavior 2025; Japan; conjoint evaluates election timing scenarios, not candidate vote choice.
Jankowski, Prokop and Tepe (2020)	Not electoral candidate choice	Hiring conjoint: German civil servants and employees evaluating job applicants for a public office position, not voters choosing among electoral candidates. Published in JPART.
Kage, Rosenbluth & Tanaka (2019), Study 2	Not electoral candidate choice	The choice is of a Liberal Democratic Party leader, not a candidate for public office. Found by the 2026-09-22 manifest audit. The paper stays in the corpus through study 1. Cleaning script in <code>cleaning_code/cleaning_archive/</code> , cleaned files in <code>clean_data/excluded_studies/</code> .
Krewson & Owens (2021)	Not electoral candidate choice	Journal of Law and Courts 2021; United States. Profiles are judicial nominees (ABA rating, prior judicial posts), not candidates in an election (<code>sen_2017</code> and <code>wu_2026</code> precedent). Found in the IRW screen (2026-10-17).
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason	Detail
Lehrer, Stockle & Juhl (2024), Study 1	Not electoral candidate choice	Party-level conjoint. The paper, p. 3: respondents are asked to choose one of the parties without an explicit none-option. The profiles are parties; the gender, age and occupation columns in our cleaned data describe each party's Spitzenkandidat rather than a candidate the respondent chose between. Mechanically eligible, so no check caught it. Its own published_estimates.csv note had read excluded from meta-analysis since ingest, and it never was. Cleaned and in the corpus until 2026-08-19; cleaning script in cleaning_code/cleaning_archive/, cleaned files in clean_data/excluded_studies/.
Sen (2017)	Not electoral candidate choice	Judicial confirmation conjoint (US Senate votes on SCOTUS nominees); not electoral candidate choice by voters.
Simon & Turnbull (2025)	Not electoral candidate choice	Meta-analysis aggregating results from other conjoint studies
Snagovsky et al. (2020)	Not electoral candidate choice	Forced choice asks which candidate 'is the most likely to care what people like you think' (external efficacy), not vote choice (Snagovsky 2019 thesis ch. 7; deposit do file 'Conjoint - External Efficacy'). Life in Australia, July 2017. Found in the IRW wave-2 ingest (2026-10-18).
Spatier (2022)	Not electoral candidate choice	Ethnic exposure in slums context; not standard candidate choice
Arceneaux & Vanderwielen (2023)	Replication data unavailable	Must be requested from authors
<i>Continued on next page</i>		

Table B.1: Studies Reviewed but Excluded (continued)

Study	Reason		Detail
Breckwoldt (2025)	Replication unavailable	data	Electoral Studies culture-war UK conjoint; binary forced-choice, 4 attributes. Data stated as available on request only — no public deposit.
Christensen, Svensson La Rosa & Gronlund (2020)	Replication unavailable	data	No public replication data available; data not archived on Dataverse, OSF, or Zenodo.
Fossati (2025)	Replication unavailable	data	Journal of Current Southeast Asian Affairs, 5 SE Asian countries; likely forced-choice, 7 attributes. No public data deposit found on Dataverse or OSF as of 2026-03-16.
Lopez Ortega and Radojevic (2025)	Replication unavailable	data	Political Behavior 2025; Germany; 2,324 voters; visual vs. text conjoint on candidate evaluation. No public data deposit found.
Magni & Reynolds (2022)	Replication unavailable	data	No accessible replication data
Sung (2022)	Replication unavailable	data	No data file available; must be requested from authors
Tsai and Wang (2025)	Replication unavailable	data	Political Studies Review Taiwan conjoint; likely forced-choice but attribute count unclear. Data stated as available on request only — no public deposit.
Weaver (2012)	Replication unavailable	data	PDF only; no data file available

Note: Covers studies reviewed from all countries, not US studies only. Every study reviewed and rejected, ordered by reason and then by study. Reasons use the same consolidation as Figure B.3. Reports removed as duplicates at the identification step are not listed, since the funnel removes them before eligibility assessment; a duplicate *sub-study* of a paper that still contributes to the corpus is listed, because its report was screened and included and the drop happened later, on the data. The table has more rows than the funnel’s 65 excluded reports, for three reasons: (i) a report contributing several rejected sub-studies appears here once per sub-study but in the funnel only once; (ii) a paper contributing at least one study to the corpus counts as an included report, so its rejected sub-study appears here but not in the exclusion box; and (iii) the two duplicate sub-studies just described are listed here, while the funnel removes neither, because both of their papers are included reports.

C Study manifest

Table C.2 lists every study in the corpus; the main text’s “Meta-reanalysis design” section refers here for the full listing. Rows are ordered by country and then alphabetically by author. For each study the table reports its country, sample type, and the number of unique respondents, binary choice tasks, and candidate profiles it contributes. The 243 studies were fielded in 52 countries. The United States contributes the most (86 studies); the next most represented are the United Kingdom (14), Germany (14), and Japan (9), and the remaining 120 studies are spread across 48 other countries. We code sample type by how respondents were recruited. Most studies (165) draw on commercial online panels that fill demographic quotas; we code a vendor panel this way even when the paper does not describe its quotas. Another 29 use probability samples, including random draws of residents or voters within purposively chosen provinces or constituencies. The rest are online convenience samples recruited through MTurk, Facebook advertisements, and similar platforms (33), samples of political elites (11), student samples (3), and in-person convenience samples (2).

Table C.2: Study Manifest

Study	Country	Sample Type	Respondents	Tasks	Profiles
Bhatia and Monroe (2024)	Afghanistan	Probability	2,485	7,455	14,910
Carnes and Lupu (2016) 1	Argentina	Probability	956	956	1,912
Gothreau and Laustsen (2026) Argentina	Argentina	Online quota	516	3,612	7,224
Laterzo (2024) 1	Argentina	Online quota	1,336	6,680	13,360
Le Foulon and Reyes-Housholder (2021) 2	Argentina	Probability	1,527	1,527	3,054
Singh (2022)	Argentina	Online quota	2,042	10,210	20,420
Kang et al. (2021)	Australia	Probability	797	3,462	6,924
Gatto and Radojevic (2025) Austria	Austria	Elite	654	5,161	10,322
Vittori, Rojon, and Pilet (2026) 1	Austria	Online quota	1,602	8,010	16,020
Vittori, Rojon, and Pilet (2026) 2	Belgium	Online quota	1,508	7,540	15,080
Vittori, Rojon, and Pilet (2026) 3	Belgium	Online quota	1,328	6,640	13,280
Walder, Bailer, and Giger (2025) 1	Belgium	Online quota	2,042	2,042	4,084
Gothreau and Laustsen (2026) Brazil	Brazil	Online quota	505	3,535	7,070
Justesen, Koob, and Smid (2025)	Brazil	Online quota	1,420	4,260	8,520
Laterzo (2024) 2	Brazil	Online quota	1,356	6,780	13,560
Smith and Boas (2024) 1	Brazil	Online convenience	1,817	5,451	10,902
Vittori, Rojon, and Pilet (2026) 4	Bulgaria	Online quota	1,601	8,005	16,010
Gidengil, Stolle and Bergeron-Boutin (2022) 1	Canada	Online quota	1,417	8,444	16,888
Gidengil, Stolle and Bergeron-Boutin (2022) 2	Canada	Online quota	1,413	8,383	16,766
Gidengil, Stolle and Bergeron-Boutin (2022b) 2	Canada	Online quota	1,817	14,456	28,912
Islam and Loewen (2026) 1	Canada	Online quota	2,165	12,990	25,980
Loewen and Rheault (2021) 2	Canada	Online quota	779	1,558	3,116
Sevi (2026) 1	Canada	Online quota	2,520	14,899	29,798
Sevi (2026) 2	Canada	Elite	85	446	892
Treger et al. (2026) 1	Canada	Online quota	1,080	3,054	6,108
Argote and Visconti (2023) 1	Chile	Online quota	3,965	19,825	39,650
Le Foulon and Reyes-Housholder (2021) 3	Chile	Probability	1,623	1,623	3,246
Segovia (2021)	Chile	Probability	1,542	6,755	13,510
Smith and Boas (2024) 2	Chile	Online convenience	3,732	11,196	22,392
Visconti (2022)	Chile	In-person convenience	201	1,467	2,934
Levy (2022)	Colombia	Online convenience	1,498	5,612	11,224
Vittori, Rojon, and Pilet (2026) 5	Czechia	Online quota	2,384	11,920	23,840
Dahl and Nyrup (2021)	Denmark	Online quota	1,621	8,105	16,210
Gothreau and Laustsen (2026) Denmark	Denmark	Online quota	508	3,556	7,112

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Littmann et al. (2026)	Denmark	Online quota	1,034	8,272	16,544
Mercea and Santos (2025) 1	Denmark	Online quota	1,001	5,005	10,010
Nyholt (2024)	Denmark	Online quota	896	3,304	6,608
Vittori, Rojon, and Pilet (2026) 6	Denmark	Online quota	1,598	7,990	15,980
Christensen et al. (2021)	Finland	Online quota	1,021	7,147	14,294
Saikkonen and Christensen (2023) 1	Finland	Online quota	1,030	6,180	12,360
Vittori, Rojon, and Pilet (2026) 7	Finland	Online quota	1,594	7,970	15,940
Coffe, Saha, and Weeks (2025) 1	France	Online quota	1,038	1,038	2,076
Gonthier and Guerra (2022)	France	Online convenience	2,743	8,229	16,458
Panel et al. (2024) 3	France	Online quota	1,625	8,125	16,250
Walder, Bailer, and Giger (2025) 2	France	Online quota	2,135	2,135	4,270
Berz and Jankowski (2022)	Germany	Elite	310	3,100	6,200
Bischof and Senninger (2026)	Germany	Online quota	14,258	42,439	84,878
Gatto and Radojevic (2025) Germany	Germany	Elite	389	3,071	6,142
Gothreau and Laustsen (2026) Germany	Germany	Online quota	1,010	7,070	14,140
Gross et al. (2026)	Germany	Online quota	2,773	8,274	16,548
Leininger and Swalve (2026)	Germany	Online quota	907	4,107	8,214
Mares, Scheve, and Toenshoff (2026)	Germany	Online quota	2,191	10,955	21,910
Mercea and Santos (2025) 2	Germany	Online quota	2,024	10,120	20,240
Neuner and Wratil (2022) 1	Germany	Online quota	2,371	11,855	23,710
Neuner and Wratil (2022) 2	Germany	Online quota	1,712	8,560	17,120
Neuner and Wratil (2022) 3	Germany	Online quota	1,714	8,570	17,140
Panel et al. (2024) 1	Germany	Online quota	1,665	8,325	16,650
Rehmert (2022)	Germany	Elite	296	1,476	2,952
Senninger and Bischof (2023)	Germany	Online convenience	993	4,965	9,930
Vittori, Rojon, and Pilet (2026) 10	Greece	Online quota	1,594	7,970	15,940
Gothreau and Laustsen (2026) Hungary	Hungary	Online quota	516	3,612	7,224
Mercea and Santos (2025) 3	Hungary	Online quota	2,051	10,255	20,510
Gothreau and Laustsen (2026) India	India	Online quota	516	3,612	7,224
Goyal (2020)	India	Probability	1,644	4,824	9,648
Rains and Wibbels (2023)	India	Probability	6,174	18,522	37,044
Thomas et al. (2024) 1	India	Probability	1,999	3,989	7,978
Thomas et al. (2024) 2	India	Probability	1,999	3,995	7,990
Gothreau and Laustsen (2026) Indonesia	Indonesia	Online quota	1,017	7,119	14,238

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Vittori, Rojon, and Pilet (2026) 11	Ireland	Online quota	1,598	7,990	15,980
Amsalem and Zoizner (2024) 2	Israel	Online quota	1,998	9,964	19,928
Franchino and Zucchini (2015)	Italy	Student	405	10,912	21,824
Martini and Olmastroni (2021)	Italy	Probability	3,096	6,192	12,384
Mercea and Santos (2025) 4	Italy	Online quota	2,101	10,505	21,010
Panel et al. (2024) 4	Italy	Online quota	1,589	7,942	15,884
Eshima and Smith (2022) 1	Japan	Online quota	2,018	20,180	40,360
Gagnon, McElwain, and Ikeda (2025) 1	Japan	Online quota	2,576	20,402	40,804
Gagnon, McElwain, and Ikeda (2025) 2	Japan	Online quota	2,580	7,689	15,378
Gothreau and Laustsen (2026) Japan	Japan	Online quota	501	3,507	7,014
Horiuchi et al. (2020)	Japan	Online quota	2,200	22,000	44,000
Kage, Rosenbluth, and Tanaka (2019) 1	Japan	Online quota	1,611	4,833	9,666
Kato, Lu, and Endo (2025) 1	Japan	Online quota	1,780	5,303	10,606
Kato, Lu, and Endo (2025) 2	Japan	Online quota	2,404	9,596	19,192
Ono and Yamada (2020)	Japan	Online quota	2,686	10,744	21,488
Gothreau and Laustsen (2026) Kenya	Kenya	Online quota	510	3,570	7,140
Cammett et al. (2022)	Lebanon	Probability	2,455	9,820	19,640
Gothreau and Laustsen (2026) Lithuania	Lithuania	Online quota	507	3,549	7,098
Ofosu et al. (2026)	Malawi	Probability	2,231	13,335	26,670
Gothreau and Laustsen (2026) Malaysia	Malaysia	Online quota	509	3,563	7,126
Gothreau and Laustsen (2026) Mexico	Mexico	Online quota	507	3,549	7,098
Greene (2024)	Mexico	Probability	1,310	1,310	2,620
Ventura, Ley, and Cantu (2024)	Mexico	Online quota	2,317	4,587	9,174
Barnett et al. (2026)	Morocco	Probability	835	3,190	6,380
Gothreau and Laustsen (2026) Morocco	Morocco	Online quota	506	3,542	7,084
Vittori, Rojon, and Pilet (2026) 16	Netherlands	Online quota	1,602	8,010	16,020
Walder, Bailer, and Giger (2025) 3	Netherlands	Online quota	1,962	1,962	3,924
Magni and Reynolds (2021) 3	New Zealand	Online quota	1,285	7,684	15,368
Gothreau and Laustsen (2026) Nigeria	Nigeria	Online quota	511	3,577	7,154
Arnesen, Duell, and Johannesson (2019) 1	Norway	Probability	1,139	3,312	6,624
Finseraas and Heim (2026)	Norway	Probability	3,171	19,026	38,052
Finseraas et al. (2026) Norway	Norway	Online quota	3,137	6,152	12,304
Vera (2025)	Paraguay	Online quota	1,933	6,866	13,732
Carter (2021)	Peru	Elite	306	1,419	2,838

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Smith and Boas (2024) 3	Peru	Online convenience	3,698	11,094	22,188
Weaver (2021)	Peru	Probability	1,037	3,788	7,576
Claassen (2026) 3	Poland	Online quota	2,010	24,120	48,240
Panel et al. (2024) 5	Poland	Online quota	1,558	7,790	15,580
Stanley (2026) 1	Poland	Online quota	3,002	68,770	137,540
Walder, Bailer, and Giger (2025) 4	Poland	Online quota	1,918	1,918	3,836
Wunsch, Jacob, and Derksen (2025)	Poland	Online quota	2,910	34,920	69,840
Gherghina and Saikkonen (2024) 1	Romania	Online quota	1,010	6,060	12,120
Gothreau and Laustsen (2026) Romania	Romania	Online quota	513	3,591	7,182
Mares and Visconti (2020) 1	Romania	In-person convenience	501	2,253	4,506
Mercea and Santos (2025) 5	Romania	Online quota	946	4,730	9,460
Gothreau and Laustsen (2026) South Africa	South Africa	Online quota	506	3,542	7,084
Claassen (2026) 2	South Korea	Online quota	2,028	24,336	48,672
Ward and Denney (2026)	South Korea	Online quota	1,983	19,830	39,660
Agerberg (2020)	Spain	Online quota	1,785	3,822	7,644
Breitenstein (2019)	Spain	Online quota	2,248	6,742	13,484
Gothreau and Laustsen (2026) Spain	Spain	Online quota	504	3,528	7,056
Lavezzolo, Ramiro, and Fernandez-Vazquez (2021)	Spain	Online quota	2,400	12,000	24,000
Panel et al. (2024) 2	Spain	Online quota	1,585	7,925	15,850
Agerberg (2024) 1	Sweden	Online convenience	4,088	4,088	8,176
Agerberg (2024) 2	Sweden	Online convenience	2,888	11,336	22,672
Agerberg (2024) 3	Sweden	Online convenience	2,939	11,590	23,180
Grahn and Thisell (2025)	Sweden	Elite	1,276	6,303	12,606
Gatto and Radojevic (2025) Switzerland	Switzerland	Elite	813	6,450	12,900
Portmann (2022)	Switzerland	Online quota	1,000	1,000	2,000
Wuest and Pontusson (2022) 1	Switzerland	Online quota	4,520	9,040	18,080
Blackman and Jackson (2021) 1	Tunisia	Probability	709	2,191	4,382
Blackman and Jackson (2021) 2	Tunisia	Online quota	574	2,870	5,740
Sasmaz (2026)	Tunisia	Elite	1,896	5,618	11,236
Gothreau and Laustsen (2026) Turkey	Turkey	Online quota	1,023	7,161	14,322
Campbell et al. (2019a) 2	UK	Online quota	1,899	9,495	18,990
Campbell et al. (2019a) 3	UK	Online quota	1,919	9,595	19,190
Campbell et al. (2019b) 1	UK	Online quota	1,719	8,595	17,190
Carnes and Lupu (2016) 2	UK	Online quota	3,783	3,783	7,566

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Coffe, Saha, and Weeks (2025) 2	UK	Online quota	1,734	1,734	3,468
Hobolt and Rodon (2020)	UK	Online quota	1,936	9,680	19,360
Hunter (2024)	UK	Online quota	1,000	3,000	6,000
Kim and Kweon (2024) 2	UK	Online convenience	982	2,946	5,892
Magni and Reynolds (2021) 2	UK	Online quota	1,120	5,549	11,098
Pedersen (2026) 1	UK	Online quota	2,231	8,924	17,848
Saha and Weeks (2022) 3	UK	Online convenience	869	4,341	8,682
Tirado Castro and Banducci (2026)	UK	Online quota	2,518	12,590	25,180
Walder, Bailer, and Giger (2025) 5	UK	Online quota	2,993	2,993	5,986
Wuest and Pontusson (2022) 3	UK	Online quota	4,241	12,722	25,444
Carlson (2015)	Uganda	Probability	751	2,237	4,474
Gothreau and Laustsen (2026) Ukraine	Ukraine	Online quota	510	3,570	7,140
Finseraas et al. (2026) United Kingdom	United Kingdom	Online quota	2,001	4,002	8,004
Mercea and Santos (2025) 6	United Kingdom	Online quota	2,224	11,120	22,240
Panel et al. (2024) 6	United Kingdom	Online quota	1,670	8,350	16,700
Reher (2025) 1	United Kingdom	Online quota	2,621	4,620	9,240
Abbot and Deveau (2024)	United States	Online quota	1,650	11,550	23,100
Albert and Costa (2025) 1	United States	Online quota	2,913	17,395	34,790
Amsalem and Zoizner (2024) 1	United States	Online quota	1,996	9,946	19,892
Arceneaux and Truex (2023)	United States	Online quota	1,883	5,553	11,106
Atkeson and Hamel (2020)	United States	Online convenience	1,500	4,500	9,000
Bakker, Schumacher, and Rooduijn (2021)	United States	Online quota	865	1,724	3,448
Bansak and Jenke (2025)	United States	Student	145	8,673	17,346
Bansak et al. (2018) 1	United States	Online convenience	4,921	147,514	295,028
Bansak et al. (2018) 2	United States	Online quota	1,617	48,407	96,814
Bansak et al. (2021) 1	United States	Online convenience	4,097	61,455	122,910
Bansak et al. (2021) 2	United States	Online quota	2,786	41,790	83,580
Carey et al. (2022) 1	United States	Elite	622	5,805	11,610
Carey et al. (2022) 2	United States	Online quota	1,000	10,000	20,000
Carey et al. (2022) 3	United States	Online quota	1,000	10,000	20,000
Carnes and Lupu (2016) 3	United States	Online quota	678	678	1,356
Castanho Silva, Neuner, and Wratil (2023)	United States	Online quota	1,485	7,393	14,786
Cho et al. (2026) 1	United States	Online quota	1,403	15,204	30,408
Cho et al. (2026) 2	United States	Online quota	1,402	15,231	30,462

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Claassen (2026) 1	United States	Online quota	2,016	24,192	48,384
Costa (2021)	United States	Online quota	1,501	9,006	18,012
Dai and Kustov (2024)	United States	Online convenience	1,004	4,016	8,032
Doherty, Dowling, and Miller (2019)	United States	Online quota	825	6,491	12,982
Driscoll and Nelson (2023)	United States	Online convenience	950	12,000	24,000
Ferrari and Smith (2025)	United States	Online convenience	2,042	12,252	24,504
Funck and McCabe (2022) 1	United States	Online quota	2,135	6,405	12,810
Gidengil, Stolle and Bergeron-Boutin (2022b) 1	United States	Online quota	1,807	14,341	28,682
Gothreau and Laustsen (2026) United States	United States	Online quota	3,129	21,903	43,806
Graham and Svolik (2020)	United States	Online quota	1,680	21,604	43,208
Green, Schaffner, and Luks (2023)	United States	Online quota	2,936	14,535	29,070
Haines and Masket (2024) 1	United States	Online quota	812	8,109	16,218
Haines and Masket (2024) 2	United States	Online quota	843	8,428	16,856
Hainmueller, Hopkins, and Yamamoto (2014)	United States	Online convenience	311	1,733	3,466
Hassell and Visalvanich (2024)	United States	Online quota	2,003	15,705	31,410
Hjortskov and Andersen (2024) 1	United States	Online convenience	1,318	6,244	12,488
Holman, Merolla, and Zechmeister (2016)	United States	Probability	1,001	1,001	2,002
Hopkins (2014)	United States	Probability	545	3,588	7,176
Islam and Loewen (2026) 2	United States	Online quota	3,612	21,672	43,344
Joo et al. (2026)	United States	Online quota	865	5,190	10,380
Kennedy et al. (2025)	United States	Online convenience	997	9,955	19,910
Kim and Kweon (2024) 1	United States	Online quota	911	2,733	5,466
Kirkland and Coppock (2018) 1	United States	Online convenience	1,203	6,004	12,008
Kirkland and Coppock (2018) 2	United States	Online quota	1,200	5,716	11,432
Kustov and Dai (2026)	United States	Online convenience	3,455	23,500	47,000
Laterzo-Tingley and Christiani (2024) 1	United States	Online quota	1,499	7,445	14,890
Laterzo-Tingley and Christiani (2024) 2	United States	Student	713	3,524	7,048
Leeper and Robison (2020)	United States	Online quota	725	3,484	6,968
Lemi (2021)	United States	Online quota	786	7,854	15,708
Loewen and Rheault (2021) 1	United States	Online quota	971	1,933	3,866
Lublin and Wright (2024)	United States	Online quota	2,314	10,556	21,112
Madrid et al. (2022)	United States	Online convenience	508	5,080	10,160
Magni and Reynolds (2021) 1	United States	Online quota	1,827	9,081	18,162
Manento and Testa (2022) 1	United States	Online convenience	1,195	5,975	11,950

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Mikkelsen (2025) 1	United States	Online quota	469	469	938
Mikkelsen (2025) 2	United States	Online quota	1,999	1,999	3,998
Mikkelsen (2025) 3	United States	Online quota	272	272	544
Mikkelsen (2025) 4	United States	Online quota	153	153	306
Mikkelsen (2025) 5	United States	Online quota	1,936	1,936	3,872
Mikkelsen (2025) 6	United States	Probability	542	542	1,084
Miller and Ziegler (2024)	United States	Online convenience	2,216	11,396	22,792
Mo (2015)	United States	Online quota	407	815	1,630
Mummolo, Peterson, and Westwood (2021) 1	United States	Online quota	3,074	21,518	43,036
Mummolo, Peterson, and Westwood (2021) 2	United States	Online quota	1,554	10,751	21,502
Noble and Carlson (2025)	United States	Online convenience	690	6,896	13,792
Ono and Burden (2019)	United States	Online quota	1,583	15,830	31,660
Ornstein et al. (2025) 1	United States	Online quota	1,308	6,540	13,080
Pedersen (2026) 2	United States	Online quota	2,185	8,740	17,480
Peterson (2017) 1	United States	Online quota	494	618	1,236
Peterson (2017) 5	United States	Online quota	528	627	1,254
Peterson and Simonovits (2018)	United States	Online quota	1,050	4,195	8,390
Reher (2025) 2	United States	Online quota	2,654	4,770	9,540
Robinson (2023)	United States	Online convenience	195	935	1,870
Saha and Weeks (2022) 1	United States	Online convenience	550	1,634	3,268
Saha and Weeks (2022) 2	United States	Online convenience	496	2,435	4,870
Saha and Weeks (2022) 4	United States	Online convenience	850	4,250	8,500
Saha and Weeks (2022) 5	United States	Online quota	1,249	3,740	7,480
Sances (2018)	United States	Online convenience	1,110	2,177	4,354
Schroeder and Campos (2026)	United States	Online convenience	599	2,381	4,762
Sevi et al. (2025)	United States	Online quota	2,102	14,337	28,674
Sevi, Crabtree, and Blais (2024)	United States	Online quota	1,971	13,701	27,402
Short et al. (2026)	United States	Online convenience	432	1,291	2,582
Teele, Kalla, and Rosenbluth (2018) 1	United States	Probability	2,105	6,225	12,450
Teele, Kalla, and Rosenbluth (2018) 2	United States	Elite	1,866	5,417	10,834
Vecchiato and Munger (2025)	United States	Online quota	500	5,000	10,000
Wakefield (2025)	United States	Online quota	810	4,833	9,666
Wood (2023)	United States	Online quota	1,487	8,882	17,764
Wuest and Pontusson (2022) 2	United States	Online quota	4,211	12,633	25,266

Continued on next page

Table C.2: Study Manifest (continued)

Study	Country	Sample Type	Respondents	Tasks	Profiles
Le Foulon and Reyes-Housholder (2021) 1	Uruguay	Probability	1,514	1,514	3,028
Malesky and Schuler (2020) 1	Vietnam	Probability	12,916	12,916	25,832
Malesky and Schuler (2020) 2	Vietnam	Probability	11,743	11,743	23,486
Malesky and Schuler (2020) 3	Vietnam	Probability	10,620	10,620	21,240
Schuler (2024)	Vietnam	Probability	5,892	5,892	11,784

Note: All 243 conjoint experiments included in the meta-reanalysis, from all countries rather than US studies only. Respondents is the number of unique survey respondents; Tasks is the number of unique binary choice tasks; Profiles is the total number of candidate profiles evaluated. Papers that contribute several studies appear as several rows.

D Generalizability

The tables in this section support the main text’s claim that the three headline patterns generalize across contexts even as the exact numerical estimates vary from study to study.

Sensitivity to the magnitude threshold

Table D.3 repeats the magnitude generalizability calculation using ± 5 and ± 15 percentage point tolerance bands. Tightening the window predictably lowers the affirmed fraction and loosening it raises the fraction, but the qualitative pattern remains the same: demographic-identity attributes such as gender often exhibit weak sign generalizability while still clustering around a near-indifferent mean, whereas political matchups retain stronger directional consistency but much lower stability in exact magnitude.

Table D.3: Magnitude generalizability sensitivity to the tolerance threshold

Attribute	k	Sign gen.	Mag. gen. $\pm 5\text{pp}$	Mag. gen. $\pm 10\text{pp}$	Mag. gen. $\pm 15\text{pp}$
Demographics					
Asian vs Hispanic	35	0.09	0.09	0.43	0.77
Black vs Hispanic	46	0.04	0.13	0.61	0.80
Christian vs Muslim	9	1.00	0.11	1.00	1.00
Christian vs None	22	0.09	0.32	0.68	0.91
Majority vs Minority	106	0.33	0.21	0.59	0.80
Middle vs Old	95	0.63	0.14	0.58	0.85
Muslim vs None	9	0.89	0.44	0.56	0.78
White vs Asian	38	0.05	0.08	0.61	0.82
White vs Black	61	0.10	0.20	0.64	0.87
White vs Hispanic	49	0.02	0.20	0.73	0.88
Women vs Men	218	0.34	0.40	0.90	0.96
Young vs Middle	133	0.11	0.32	0.80	0.92
Young vs Old	114	0.61	0.18	0.53	0.83
Experience					
Business/Managerial vs Professional	110	0.12	0.04	0.53	0.85
College Degree vs None	55	0.67	0.27	0.75	0.96
Educator vs Professional	75	0.03	0.09	0.40	0.79
Farmer vs Professional	35	0.03	0.09	0.37	0.74
Military vs Professional	14	0.07	0.29	0.64	0.79
No vs Any Political Experience	120	0.62	0.15	0.57	0.92
Working Class vs Professional	35	0.14	0.20	0.60	0.89
Politics					
Conservative vs Liberal Issue	95	0.55	0.13	0.37	0.55
Positions					
Conservative vs Mixed Issue Positions	76	0.49	0.04	0.53	0.83
Democrat vs Independent	8	0.12	0.12	0.88	1.00
In-party vs Out-party	79	0.99	0.11	0.46	0.80
Liberal vs Mixed Issue Positions	76	0.37	0.11	0.53	0.74
Republican vs Democrat	49	0.45	0.31	0.73	0.96
Republican vs Independent	8	1.00	0.12	0.75	1.00

Estimates cover conjoints from all countries, not US studies only. Sign gen. is the Egami–Hartman sequential-test fraction of study-level estimates affirmed to lie on the same side of 50% as $\hat{\mu}$. Magnitude fractions use the same procedure but test whether estimates lie within a symmetric band around $\hat{\mu}$. The three columns report bands of ± 5 , ± 10 , and ± 15 percentage points.

US sign generalizability

Tables D.4 and D.5 repeat the sign and magnitude generalizability calculations within the US studies, separately for Republican and for Democratic respondents. Splitting by respondent party separates the two ways an attribute can fail to generalize: a matchup can point the same way in every study while respondents of the two parties disagree about which way that is, and a matchup on which the parties agree can still vary widely in magnitude from study to study.

Table D.4: US sign and magnitude generalizability, Republican respondents

Attribute	k	μ (SE)	τ (SE)	Q_p	Sign gen.	Mag. gen.
Demographics						
Women vs Men	62	0.49 (0.00)	0.01 (0.01)	0.007	0.02	0.76
Young vs Old	31	0.53 (0.01)	0.03 (0.02)	<0.001	0.19	0.42
Young vs Middle	35	0.49 (0.00)	0.01 (0.01)	0.066	0.03	0.57
Middle vs Old	29	0.56 (0.01)	0.02 (0.01)	0.001	0.48	0.48
Majority vs Minority	51	0.52 (0.00)	0.03 (0.01)	<0.001	0.06	0.63
White vs Black	50	0.52 (0.01)	0.03 (0.02)	<0.001	0.06	0.32
White vs Hispanic	44	0.53 (0.01)	0.03 (0.02)	<0.001	0.09	0.34
White vs Asian	32	0.52 (0.01)	0.04 (0.03)	<0.001	0.03	0.22
Asian vs Hispanic	30	0.54 (0.01)	0.04 (0.03)	<0.001	0.10	0.20
Black vs Hispanic	42	0.50 (0.00)	0.01 (0.01)	0.380	0.02	0.31
Christian vs None	13	0.53 (0.01)	0.02 (0.02)	0.060	0.08	0.62
Christian vs Muslim	3	0.74 (0.02)	0.00 (0.04)	0.269	1.00	0.67
Muslim vs None	3	0.34 (0.02)	0.00 (0.07)	0.745	0.33	0.33
Experience						
College Degree vs None	12	0.59 (0.02)	0.05 (0.04)	<0.001	0.50	0.58
No vs Any Political Experience	32	0.43 (0.01)	0.06 (0.03)	<0.001	0.59	0.47
Business/Managerial vs Professional	31	0.53 (0.01)	0.03 (0.02)	<0.001	0.10	0.29
Educator vs Professional	27	0.51 (0.01)	0.05 (0.04)	0.004	0.04	0.19
Farmer vs Professional	18	0.51 (0.01)	0.02 (0.02)	0.326	0.06	0.28
Military vs Professional	8	0.53 (0.03)	0.05 (0.05)	0.058	0.12	0.12
Working Class vs Professional	13	0.49 (0.02)	0.05 (0.04)	<0.001	0.23	0.31
Politics						
Conservative vs Liberal Issue	38	0.66 (0.02)	0.12 (0.06)	<0.001	0.58	0.16
Positions						
Conservative vs Mixed Issue Positions	36	0.60 (0.01)	0.08 (0.04)	<0.001	0.47	0.14
Liberal vs Mixed Issue Positions	36	0.36 (0.02)	0.10 (0.05)	<0.001	0.58	0.14
In-party vs Out-party	46	0.72 (0.01)	0.09 (0.04)	<0.001	1.00	0.37
Republican vs Democrat	46	0.72 (0.01)	0.09 (0.04)	<0.001	1.00	0.37
Republican vs Independent	7	0.60 (0.02)	0.04 (0.04)	0.110	0.29	0.14
Democrat vs Independent	7	0.31 (0.03)	0.07 (0.06)	<0.001	0.86	0.00

Each row summarizes one matchup among US studies, for Republican respondents. μ is the meta-analytic mean AFCP and τ is the between-study standard deviation (REML). Q_p is the p -value from Cochran's Q test of residual homogeneity. Sign gen. is the fraction of study-level estimates affirmed to lie on the same side of 50% as $\hat{\mu}$ by the Egami-Hartman sequential test. Mag. gen. is the fraction affirmed to lie within ± 10 percentage points of $\hat{\mu}$ by the same procedure. Standard errors in parentheses.

Table D.5: US sign and magnitude generalizability, Democratic respondents

Attribute	k	μ (SE)	τ (SE)	Q_p	Sign gen.	Mag. gen.
Demographics						
Women vs Men	66	0.54 (0.01)	0.04 (0.02)	<0.001	0.41	0.65
Young vs Old	31	0.56 (0.01)	0.04 (0.02)	<0.001	0.35	0.48
Young vs Middle	38	0.50 (0.01)	0.03 (0.02)	<0.001	0.03	0.55
Middle vs Old	29	0.56 (0.01)	0.03 (0.02)	<0.001	0.52	0.62
Majority vs Minority	56	0.47 (0.01)	0.04 (0.02)	<0.001	0.21	0.71
White vs Black	55	0.47 (0.01)	0.04 (0.02)	<0.001	0.16	0.45
White vs Hispanic	46	0.49 (0.00)	0.02 (0.01)	<0.001	0.04	0.50
White vs Asian	34	0.48 (0.01)	0.05 (0.03)	<0.001	0.09	0.35
Asian vs Hispanic	33	0.51 (0.01)	0.04 (0.03)	<0.001	0.09	0.33
Black vs Hispanic	44	0.52 (0.01)	0.03 (0.02)	<0.001	0.05	0.43
Christian vs None	13	0.48 (0.01)	0.01 (0.01)	0.135	0.23	0.69
Christian vs Muslim	3	0.57 (0.02)	0.03 (0.04)	0.218	0.33	1.00
Muslim vs None	3	0.39 (0.06)	0.10 (0.11)	0.016	0.67	0.33
Experience						
College Degree vs None	12	0.59 (0.02)	0.06 (0.04)	<0.001	0.67	0.50
No vs Any Political Experience	35	0.41 (0.01)	0.06 (0.03)	<0.001	0.66	0.46
Business/Managerial vs Professional	34	0.49 (0.01)	0.03 (0.02)	<0.001	0.06	0.26
Educator vs Professional	29	0.52 (0.01)	0.04 (0.03)	<0.001	0.10	0.14
Farmer vs Professional	18	0.50 (0.01)	0.00 (0.01)	0.046	0.00	0.28
Military vs Professional	8	0.49 (0.01)	0.00 (0.03)	0.707	0.00	0.50
Working Class vs Professional	13	0.49 (0.01)	0.01 (0.01)	0.297	0.08	0.46
Politics						
Conservative vs Liberal Issue Positions	39	0.29 (0.02)	0.10 (0.05)	<0.001	0.72	0.13
Conservative vs Mixed Issue Positions	36	0.32 (0.02)	0.10 (0.05)	<0.001	0.86	0.19
Liberal vs Mixed Issue Positions	36	0.65 (0.02)	0.10 (0.05)	<0.001	0.69	0.25
In-party vs Out-party	46	0.73 (0.01)	0.09 (0.04)	<0.001	0.98	0.35
Republican vs Democrat	46	0.27 (0.01)	0.09 (0.04)	<0.001	0.98	0.35
Republican vs Independent	7	0.25 (0.03)	0.07 (0.06)	<0.001	1.00	0.29
Democrat vs Independent	7	0.60 (0.01)	0.02 (0.03)	0.229	0.57	0.43

Each row summarizes one matchup among US studies, for Democratic respondents. μ is the meta-analytic mean AFCP and τ is the between-study standard deviation (REML). Q_p is the p -value from Cochran's Q test of residual homogeneity. Sign gen. is the fraction of study-level estimates affirmed to lie on the same side of 50% as $\hat{\mu}$ by the Egami-Hartman sequential test. Mag. gen. is the fraction affirmed to lie within ± 10 percentage points of $\hat{\mu}$ by the same procedure. Standard errors in parentheses.

E Validation against published estimates

Figure E.4 plots our re-estimated AFCPs against published estimates translated to the AFCP scale, separately for comparisons the original article reports numerically and those it displays only in a coefficient plot. Points cluster along the 45-degree line of perfect agreement in both panels, more tightly for numerically reported estimates, confirming that our cleaning and estimation procedures recover the published findings. Not every figure-only comparison can be read off its plot as a single number: where an article displays a multi-level attribute level by level and our matchup pools those levels, the display supports the comparison without stating a value we can translate onto the AFCP scale. We therefore validate against all 341 comparisons an article states numerically and against the 76 figure-only comparisons for which a single published value can be recovered. The main text’s “Meta-reanalysis design” section summarizes this validation and points here for the detail. Table E.6 summarizes the agreement. The full comparison, one row per (study, matchup) pair, is included in the replication archive as `validation_published_estimates_full.csv`.

Reporting	Comparisons	Studies	Median absolute discrepancy (pp)	Within 2 pp (%)	Within 5 pp (%)
Article prints a number	341	107	0.5	86	97
Figure only (read from plot)	76	33	0.6	86	99

Table E.6: Agreement between our re-estimated study-level AFCPs and published estimates, by how the original article reports the contrast

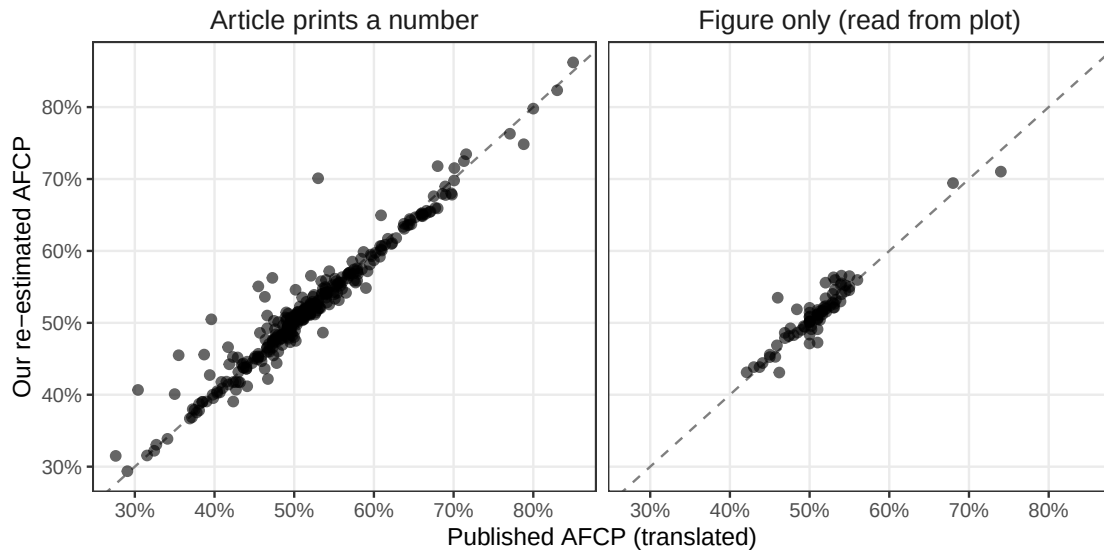


Figure E.4: Re-estimated AFCPs versus published AFCPs (translated to the AFCP scale)
Note: Estimates cover conjoints from all countries, not US studies only. Each point is one (study, matchup) pair for which a published value exists and can be translated to the AFCP scale ($n = 417$). Panels split comparisons by how the original article reports the contrast: as a printed numeric estimate (left) or only in a coefficient plot read off the figure (right). The dashed line marks perfect agreement ($y = x$).

F Additional regression tables

F.1 Heterogeneity by respondent partisanship

In the main text, we present figures that show how AFCPs may differ by respondent partisanship. Table F.7 presents formal tests of the differences in AFCPs. We see substantively very large (and statistically significant) differences for the politics variables, substantively small and occasionally significant differences for demographic variables, and only occasionally significant differences for experience variables.

Table F.7: Republican versus Democratic preferences for candidate attributes

	Baseline AFCP	Dem. – Rep.		
	Republicans	Difference	k	p_Q
Politics				
Republican vs Democrat	0.724*	-0.452*	92	<0.001
	(0.014)	(0.019)		
Republican vs Independent	0.597*	-0.348*	14	<0.001
	(0.029)	(0.038)		
Democrat vs Independent	0.312*	0.284*	14	<0.001
	(0.025)	(0.034)		
In-party vs Out-party	0.724*	0.004	92	<0.001
	(0.014)	(0.019)		
Conservative vs Liberal Issue Positions	0.656*	-0.369*	77	<0.001
	(0.020)	(0.028)		
Conservative vs Mixed Issue Positions	0.598*	-0.276*	72	<0.001
	(0.017)	(0.024)		
Liberal vs Mixed Issue Positions	0.359*	0.296*	72	<0.001
	(0.018)	(0.025)		
Demographics				
Women vs Men	0.490*	0.051*	128	<0.001
	(0.004)	(0.006)		
Young vs Old	0.532*	0.027*	62	<0.001
	(0.008)	(0.012)		
Young vs Middle	0.493	0.009	73	<0.001

Table F.7: Republican versus Democratic preferences for candidate attributes (*continued*)

	Republicans	Difference	k	p_Q
	(0.005)	(0.007)		
Middle vs Old	0.558*	0.004	58	<0.001
	(0.006)	(0.009)		
Majority vs Minority	0.519*	-0.047*	107	<0.001
	(0.006)	(0.008)		
White vs Black	0.522*	-0.055*	105	<0.001
	(0.007)	(0.009)		
White vs Hispanic	0.530*	-0.043*	90	<0.001
	(0.006)	(0.008)		
White vs Asian	0.516	-0.034*	66	<0.001
	(0.010)	(0.014)		
Black vs Hispanic	0.505	0.016	86	<0.001
	(0.007)	(0.010)		
Asian vs Hispanic	0.541*	-0.033*	63	<0.001
	(0.011)	(0.015)		
Christian vs None	0.534*	-0.050*	26	0.036
	(0.008)	(0.010)		
Christian vs Muslim	0.738*	-0.165*	6	0.225
	(0.026)	(0.035)		
Muslim vs None	0.355*	0.037	6	0.065
	(0.056)	(0.075)		
Experience				
No vs Any Political Experience	0.429*	-0.017	67	<0.001
	(0.011)	(0.016)		
College Degree vs None	0.586*	0.003	24	<0.001
	(0.018)	(0.025)		
Working Class vs Professional	0.497	-0.012	26	<0.001
	(0.012)	(0.016)		
Business/Managerial vs Professional	0.530*	-0.039*	65	<0.001
	(0.009)	(0.012)		
Educator vs Professional	0.509	0.008	56	<0.001

Table F.7: Republican versus Democratic preferences for candidate attributes (*continued*)

	Republicans	Difference	k	p_Q
	(0.013)	(0.017)		
Farmer vs Professional	0.504	-0.001	36	0.068
	(0.008)	(0.011)		
Military vs Professional	0.537*	-0.045	16	0.195
	(0.018)	(0.023)		

Estimates are restricted to US studies. Random-effects meta-regression (REML). Baseline AFCP is estimated candidate A win share among Republican respondents, with stars indicating departure from the null of 50%. Difference tests whether the AFCP differs between Democrat and Republican respondents. p_Q is the p-value from Cochran's Q heterogeneity test. Standard errors in parentheses. * $p < 0.05$.

F.2 Heterogeneity by presence of policy information

In the main text we argue that attribute effects are smaller among studies that display policy information. Table F.8 reports the policy-information regressions that underlie that claim. A study counts as displaying policy information when the conjoint shows candidates' stated positions on matters of government action; attributes that give only an ideological label, that name the issue a candidate would prioritize without stating a position, or that convey character or values without reference to policy, such as a candidate's ethical judgments, are not coded as policy information.

Table F.8: Descriptive contrasts across studies with and without policy information

	Baseline AFCP	Difference		
	No policy	Policy incl.	k	p_Q
Politics				
Republican vs Democrat	0.442*	0.041*	49	<0.001
	(0.009)	(0.012)		
Republican vs Independent	0.359*	0.049	8	0.017
	(0.018)	(0.026)		

Table F.8: Descriptive contrasts across studies with and without policy information (*continued*)

	No policy	Policy incl.	k	p_Q
Democrat vs Independent	0.453* (0.013)	0.052* (0.019)	8	0.251
In-party vs Out-party	0.781* (0.010)	-0.104* (0.014)	79	<0.001
Demographics				
Women vs Men	0.536* (0.004)	-0.013* (0.005)	218	<0.001
Young vs Old	0.588* (0.006)	-0.028* (0.010)	114	<0.001
Young vs Middle	0.508* (0.004)	0.007 (0.006)	133	<0.001
Middle vs Old	0.597* (0.006)	-0.047* (0.010)	95	<0.001
Majority vs Minority	0.542* (0.008)	-0.041* (0.013)	106	<0.001
White vs Black	0.477* (0.008)	0.011 (0.011)	61	<0.001
White vs Hispanic	0.503 (0.007)	-0.002 (0.008)	49	<0.001
White vs Asian	0.486 (0.011)	0.025 (0.015)	38	<0.001
Black vs Hispanic	0.509 (0.009)	0.012 (0.012)	46	<0.001
Asian vs Hispanic	0.546* (0.013)	-0.040* (0.018)	35	<0.001
Christian vs None	0.485 (0.012)	0.007 (0.017)	22	<0.001
Christian vs Muslim	0.615* (0.014)	-0.035 (0.028)	9	<0.001

Table F.8: Descriptive contrasts across studies with and without policy information (*continued*)

	No policy	Policy incl.	k	p_Q
Muslim vs None	0.368* (0.014)	0.009 (0.028)	9	<0.001
Experience				
No vs Any Political Experience	0.407* (0.006)	0.061* (0.010)	120	<0.001
College Degree vs None	0.570* (0.010)	-0.011 (0.016)	55	<0.001
Working Class vs Professional	0.476* (0.011)	0.028 (0.017)	35	<0.001
Business/Managerial vs Professional	0.488* (0.005)	-0.010 (0.007)	110	<0.001
Educator vs Professional	0.493 (0.006)	0.011 (0.011)	75	<0.001
Farmer vs Professional	0.507 (0.008)	0.003 (0.012)	35	<0.001
Military vs Professional	0.487 (0.015)	0.001 (0.018)	14	0.019

Estimates cover conjoints from all countries, not US studies only. Random-effects meta-regression (REML). Baseline AFCP is estimated candidate A win share in studies without policy positions shown, with stars indicating departure from the null of 50%. Difference tests whether the AFCP changes when policy information is included. p_Q is the p-value from Cochran's Q heterogeneity test. Standard errors in parentheses. $*p < 0.05$.

Table F.9 reports the results of a meta-analysis of the studies that randomize policy information at the task level. Four studies do so, and all four contribute. Study 1 of Peterson (2017) enters only the in-party matchup, which is therefore the one row where k reaches four.

Table F.9: Causal effect estimates of policy information among randomized studies

	Baseline AFCP	Difference		
	No policy	Policy shown	k	p_Q
Politics				
In-party vs Out-party	0.795* (0.028)	-0.107* (0.012)	4	0.016
Demographics				
Women vs Men	0.508 (0.007)	-0.005 (0.008)	3	0.171
Young vs Old	0.599* (0.015)	-0.067* (0.016)	3	0.799
Young vs Middle	0.503 (0.009)	0.005 (0.010)	3	0.977
Middle vs Old	0.578* (0.012)	-0.018 (0.024)	3	0.062
Majority vs Minority	0.513 (0.018)	0.001 (0.009)	3	0.543
White vs Black	0.491 (0.019)	0.022 (0.016)	3	0.088
White vs Hispanic	0.542* (0.013)	-0.035* (0.014)	3	0.890
White vs Asian	0.515 (0.030)	-0.005 (0.028)	3	0.073
Black vs Hispanic	0.539* (0.013)	-0.043* (0.014)	3	0.446
Asian vs Hispanic	0.534* (0.016)	-0.023 (0.014)	3	0.190
Christian vs None	0.493 (0.010)	0.004 (0.009)	3	0.335
Experience				
No vs Any Political Experience	0.407*	0.058*	2	0.066

Table F.9: Causal effect estimates of policy information among randomized studies (*continued*)

	No policy	Policy shown	k	p_Q
College Degree vs None	(0.022) 0.666*	(0.020) -0.085*	2	0.001
Working Class vs Professional	(0.030) 0.515	(0.019) -0.006	2	0.141
Business/Managerial vs Professional	(0.025) 0.546*	(0.023) -0.032*	3	0.501
Farmer vs Professional	(0.014) 0.518	(0.015) -0.021	3	0.417
	(0.013)	(0.014)		

Estimates are restricted to US studies. Random-effects meta-regression (REML). Restricted to the four studies that randomize whether policy information is present at the task level (Bansak et al. 2018 studies 1 and 2; Funck and McCabe 2022; Peterson 2017 study 1). Peterson 2017 study 1 contributes only to the in-party row, the one row where k reaches four: in every other matchup its tasks all fall in the no-policy condition, so it supplies no within-study contrast. Within each study, A is winner is regressed on task_policy_shown (0 = no policy attributes shown, 1 = at least one policy attribute shown). Baseline AFCP is the meta-analytic mean intercept (expected win share when policy is not shown), with stars indicating departure from the null of 50%. Difference is the meta-analytic mean slope (change in AFCP when policy is shown), with stars indicating departure from 0. Policy matchups (c vs l, etc.) are omitted because they require policy information to be shown and therefore have no variation in task_policy_shown within the relevant tasks. p-Q is the p-value from Cochran's Q heterogeneity test on the slope estimates. Standard errors in parentheses. $*p < 0.05$.

Election type and level

In the main text’s “Additional Analyses” section we argue that attribute effects are largely stable across the type and level of office a study simulates, with the age matchups in regional elections as the clearest exception. In-party preferences are larger in point estimate in local and regional elections than in national ones, but neither difference is significant. Tables F.10 and F.11 report the underlying meta-regressions. For nearly every attribute the office-type and office-level coefficients are statistically indistinguishable from zero, so the baseline estimates do not depend on the office context. The in-party advantage is larger in local and regional contests, but neither difference is significant at the 5% level.

Table F.10: Effect of election type on candidate attribute preferences

	Baseline AFCP	Difference		
	Legislative	Exec. — Leg.	<i>k</i>	<i>pQ</i>
Politics				
Republican vs Democrat	0.475* (0.008)	-0.014 (0.014)	47	<0.001
Republican vs Independent	0.395* (0.021)	-0.015 (0.029)	10	0.002
Democrat vs Independent	0.494 (0.015)	-0.041 (0.022)	10	0.085
In-party vs Out-party	0.724* (0.012)	0.007 (0.024)	69	<0.001
Conservative vs Liberal Issue Positions	0.414* (0.013)	0.037 (0.021)	91	<0.001
Conservative vs Mixed Issue Positions	0.425* (0.010)	0.043* (0.020)	72	<0.001
Liberal vs Mixed Issue Positions	0.545* (0.012)	-0.042 (0.022)	72	<0.001
Demographics				
Women vs Men	0.534* (0.004)	-0.010 (0.006)	202	<0.001

Table F.10: Effect of election type on candidate attribute preferences (*continued*)

	Legislative	Exec. — Leg.	k	p_Q
Young vs Old	0.590* (0.007)	-0.038* (0.011)	104	<0.001
Young vs Middle	0.514* (0.003)	-0.027* (0.008)	123	<0.001
Middle vs Old	0.593* (0.006)	-0.039* (0.013)	86	<0.001
Majority vs Minority	0.535* (0.008)	-0.026 (0.018)	94	<0.001
White vs Black	0.484* (0.005)	0.019* (0.009)	52	<0.001
White vs Hispanic	0.496 (0.005)	0.020* (0.008)	42	<0.001
White vs Asian	0.492 (0.011)	0.021 (0.017)	35	<0.001
Black vs Hispanic	0.507 (0.005)	0.001 (0.007)	35	0.470
Asian vs Hispanic	0.531* (0.011)	-0.019 (0.018)	27	<0.001
Christian vs None	0.481 (0.012)	0.016 (0.017)	22	<0.001
Christian vs Muslim	0.608* (0.014)	-0.018 (0.050)	9	<0.001
Muslim vs None	0.372* (0.012)	-0.056 (0.059)	9	<0.001
Experience				
No vs Any Political Experience	0.428* (0.006)	0.007 (0.014)	116	<0.001
College Degree vs None	0.560* (0.009)	0.033 (0.020)	54	<0.001
Working Class vs Professional	0.489	-0.012	33	<0.001

Table F.10: Effect of election type on candidate attribute preferences (*continued*)

	Legislative	Exec. – Leg.	k	p_Q
Business/Managerial vs Professional	(0.010) 0.490*	(0.021) -0.022*	101	<0.001
Educator vs Professional	(0.004) 0.494	(0.009) 0.016	76	<0.001
Farmer vs Professional	(0.005) 0.502	(0.017) 0.001	26	0.013
Military vs Professional	(0.007) 0.489	(0.013) -0.012	12	0.010
	(0.010)	(0.019)		

Estimates cover conjoints from all countries, not US studies only. Random-effects meta-regression (REML). Baseline AFCP is estimated candidate A win share in legislative election studies, with stars indicating departure from the null of 50%. Difference tests whether the AFCP differs in executive vs legislative election studies. p_Q is the p-value from Cochran's Q heterogeneity test. Standard errors in parentheses. $*p < 0.05$.

Table F.11: Effect of election level on candidate attribute preferences

	Baseline AFCP	Difference			
	National	Local – Nat.	Reg. – Nat.	k	p_Q
Politics					
Republican vs Democrat	0.473* (0.006)	-0.039 (0.025)	-0.033 (0.031)	52	<0.001
Republican vs Independent	0.410* (0.018)	-0.049 (0.026)	-0.024 (0.040)	10	0.027
Democrat vs Independent	0.509 (0.012)	-0.061* (0.018)	-0.056 (0.037)	10	0.522
In-party vs Out-party	0.721* (0.010)	0.023 (0.030)	0.097 (0.050)	76	<0.001
Conservative vs Liberal Issue Positions	0.426* (0.010)	0.090* (0.044)	0.004 (0.044)	98	<0.001
Conservative vs Mixed Issue Positions	0.433* (0.009)	0.054 (0.035)	0.015 (0.031)	79	<0.001
Liberal vs Mixed Issue Positions	0.538* (0.010)	-0.063 (0.038)	-0.026 (0.034)	79	<0.001
Demographics					
Women vs Men	0.530* (0.003)	-0.003 (0.009)	0.005 (0.010)	224	<0.001
Young vs Old	0.576* (0.005)	-0.013 (0.023)	-0.070* (0.022)	119	<0.001
Young vs Middle	0.512* (0.003)	-0.011 (0.010)	-0.046* (0.014)	137	<0.001
Middle vs Old	0.582* (0.006)	-0.023 (0.023)	-0.022 (0.022)	100	<0.001
Majority vs Minority	0.525* (0.007)	-0.022 (0.023)	0.046 (0.030)	109	<0.001

Table F.11: Effect of election level on candidate attribute preferences (*continued*)

	National	Local – Nat.	Reg. – Nat.	k	p_Q
White vs Black	0.483* (0.006)	0.005 (0.015)	0.036 (0.031)	64	<0.001
White vs Hispanic	0.502 (0.004)	0.009 (0.014)	0.003 (0.022)	52	<0.001
White vs Asian	0.500 (0.008)	-0.011 (0.024)	0.067 (0.055)	41	<0.001
Black vs Hispanic	0.518* (0.007)	-0.007 (0.019)	-0.006 (0.036)	44	<0.001
Asian vs Hispanic	0.533* (0.011)	-0.047 (0.029)		33	<0.001
Christian vs None	0.488 (0.009)	0.015 (0.029)	0.012 (0.044)	24	<0.001
Experience					
No vs Any Political Experience	0.435* (0.006)	-0.017 (0.020)	-0.039 (0.023)	124	<0.001
College Degree vs None	0.568* (0.008)	-0.021 (0.060)	0.010 (0.060)	53	<0.001
Working Class vs Professional	0.488 (0.009)	0.006 (0.024)		35	<0.001
Business/Managerial vs Professional	0.482* (0.004)	0.023 (0.016)	0.011 (0.020)	110	<0.001
Educator vs Professional	0.489* (0.005)	0.067* (0.014)	0.017 (0.017)	80	<0.001
Farmer vs Professional	0.509 (0.006)		-0.009 (0.024)	35	<0.001

Random-effects meta-regression (REML). Baseline AFCP is estimated candidate A win share in national election studies, with stars indicating departure from the null of 50%. Differences test whether the AFCP differs in local or regional vs national election studies. p_Q is the p-value from Cochran's Q heterogeneity test. Standard errors in parentheses. * $p < 0.05$.

G Attenuation in the number of attributes

This section backs the main text’s “Information as a Moderator of Candidate Evaluation” argument that the attenuation of party cues under policy information reflects policy content specifically, not a mechanical artifact of longer candidate profiles.

Violations of the independence of irrelevant alternatives imply that AFCPs should weakly attenuate toward indifference as the number of randomized attributes grows, whatever the content of those attributes (Abramson, Koçak, and Magazinnik, 2022). Table G.12 tests this prediction. The pooled models regress each study-matchup estimate’s oriented distance from 0.5 on the study’s attribute count, with matchup fixed effects, study and study-by-matchup random effects, and cluster-robust standard errors by study; orientation is leave-one-study-out, so no estimate’s sign depends on its own value. Pooled attenuation is small and statistically indistinguishable from zero. The per-matchup regressions show that attenuation concentrates where the cue is strongest: the in-party advantage, whose baseline sits 22.7 points from indifference, shrinks by 1.1 percentage points per additional attribute ($p = 0.022$). The penalty for political inexperience, whose baseline sits only 6.9 points from indifference, moves by 0.2 points per attribute, which we cannot distinguish from zero ($p = 0.49$). Three further slopes are distinguishable from zero. The Christian-versus-none effect moves further from indifference as profiles lengthen, which is not attenuation. The Democrat-versus-Independent and working-class effects do move toward indifference, but both start close to it (baselines of 0.474 and 0.487), so the slopes of 2.2 and 0.6 points per attribute act on gaps of under three points. Neither pattern is the across-the-board dilution the mechanism predicts.

Table G.13 enters policy content and attribute count as competing moderators for these two matchups. For the in-party matchup, policy content does most of the work: displaying policy positions reduces the AFCP by 9.9 percentage points conditional on attribute count, most of its unconditional effect of 10.4 points, while each additional attribute of any kind contributes a further 0.6 points, which we cannot distinguish from zero ($p = 0.13$). For experience, policy content does the work alone: conditional on attribute count, displaying policy positions attenuates the penalty by 6.1 percentage points ($p < 0.001$), while each additional attribute contributes 0.1 points ($p = 0.63$). Policy content outweighs attribute count by roughly ten to one for the party cue and by far more for experience.

Table G.12: AFCP attenuation in the number of randomized attributes

	Baseline AFCP	Slope per attribute	<i>k</i>
Pooled (leave-one-study-out oriented magnitude)			
All matchups, linear (pooled, matchup FE)		-0.001 (0.001)	1660
All matchups, log attribute count (pooled, matchup FE)		-0.003 (0.008)	1660
By matchup (AFCP on attribute count)			
In-party vs Out-party	0.727	-0.011* (0.005)	71
Republican vs Independent	0.378	0.014 (0.016)	8
Middle vs Old	0.579	-0.005 (0.004)	87
Young vs Old	0.578	-0.004 (0.004)	106
Conservative vs Liberal Issue Positions	0.427	-0.006 (0.006)	90
No vs Any Political Experience	0.431	0.002 (0.003)	116
Conservative vs Mixed Issue Positions	0.436	0.002 (0.004)	71
College Degree vs None	0.561	0.000 (0.004)	51
Asian vs Hispanic	0.534	0.008 (0.005)	29
Republican vs Democrat	0.467	0.003 (0.003)	41
Women vs Men	0.531	-0.002	212

Table G.12: AFCP attenuation in the number of randomized attributes (*continued*)

	Baseline AFCP	Slope per attribute	<i>k</i>
Liberal vs Mixed Issue Positions	0.530	(0.002) 0.008	71
Democrat vs Independent	0.474	(0.005) 0.022*	8
Majority vs Minority	0.525	(0.010) 0.001	100
Farmer vs Professional	0.522	(0.003) 0.001	29
Black vs Hispanic	0.518	(0.005) -0.005	40
White vs Black	0.482	(0.003) 0.004	55
Business/Managerial vs Professional	0.485	(0.003) -0.002	104
Christian vs None	0.485	(0.002) -0.010*	16
Working Class vs Professional	0.487	(0.002) 0.006	31
Military vs Professional	0.487	(0.003) 0.016*	12
Educator vs Professional	0.491	(0.007) -0.004	73
Young vs Middle	0.508	(0.004) -0.000	125
White vs Asian	0.494	(0.002) 0.005	32
White vs Hispanic	0.504	(0.004) 0.003	43
		(0.002)	

Note: Estimates cover conjoints from all countries, not US studies only. Each study contributes one estimate per matchup. The nine studies whose attribute count `meta_data/conjoint_studies.xlsx` leaves blank are excluded, since the regressor is undefined for them; for most the count is blank because the design randomizes or varies which attributes appear, and for the rest it is simply not recoverable from the report. Pooled models regress the leave-one-study-out oriented distance from 0.5 on attribute count with matchup fixed effects, study and study-matchup random effects, and CR2 cluster-robust standard errors by study. Per-matchup models regress the AFCP on attribute count within matchups with at least 8 studies (random effects). Standard errors in parentheses. * $p < 0.05$.

Table G.13: Policy content versus attribute count as competing moderators of AFCPs

	In-party vs Out-party	No vs Any Political Experience
Policy shown (unconditional)	-0.104*	0.061*
	(0.016)	(0.010)
Attribute count (unconditional)	-0.011*	0.002
	(0.005)	(0.003)
Policy shown (conditional)	-0.099*	0.061*
	(0.016)	(0.010)
Attribute count (conditional)	-0.006	0.001
	(0.004)	(0.003)
k	71	116

Note: Random-effects meta-regressions of study-level AFCPs on whether policy positions were shown and the number of randomized attributes, entered separately (unconditional) and together (conditional), within each matchup. The nine studies whose attribute count is not recorded are excluded, since the regressor is undefined for them. Standard errors in parentheses. * $p < 0.05$.

H Recoverable no-choice rates

In a small subset of studies, the raw data preserve abstention or no-choice behavior rather than only the forced-choice winner. The main text’s “Additional Analyses” section flags this audit and points here. Table H.14 summarizes the recoverable rates in those designs. The observed no-choice rates are often nontrivial when the option is available, but the archival support for these outcomes is too sparse and inconsistent to justify a pooled abstention meta-analysis for the full corpus. See Visconti and Yang (2025) for the methodological considerations that attend conjoints with an explicit abstention option.

Table H.14: Recoverable no-choice rates in studies that permit abstention

Study	Country	Office	No-choice option available	No-choice rate
Carnes and Lupu (2016) study 1	Argentina	Mayor	All ballots	11.5%
Carnes and Lupu (2016) study 2	United Kingdom	Local political office	All ballots	31.9%
Carnes and Lupu (2016) study 3	United States	State legislature / Congress	All ballots	32.2%
Snagovsky et al. (2020) study 1	Australia	Federal parliament	All ballots	26.6%
Martin and Blinder (2021) study 1	United Kingdom	Candidate-choice ballots	50.1% of ballots	20.3% overall (40.5% conditional on no-choice option)

Note: The main stacked analysis file is built around forced-choice outcomes, so these rates are recoverable only in a small subset of studies where the raw data preserve abstention or no-choice behavior. Carnes and Lupu (2016) and Snagovsky et al. (2020) allow direct recovery from raw ballot outcomes. Martin and Blinder (2021) include a no-choice option on only some ballots; its abstention rate is reconstructed from paired candidate rows because the archived file does not include respondent or task identifiers.

I Satisficing correction subset robustness

Clayton et al. (2026) show that some conjoint respondents answer haphazardly rather than reading the profiles, and that the resulting measurement error attenuates estimated effects toward indifference. Their correction models each observed choice as a mixture of a real choice and a coin flip and recovers the mixing rate from repeated identical tasks presented to the same respondent. We adapt this correction to the AFCEP: Clayton et al. (2026) develop it for the AMCE, whereas we estimate a study-specific swapping-error parameter from repeated identical tasks and use it to reinflate each study’s AFCEP away from 0.5, propagating the added uncertainty into the confidence intervals. The correction is feasible only where the raw data preserve repeated identical tasks, so we report it for that subset rather than the full corpus. This analysis backs the main text’s treatment of the satisficing-masking tradeoff in the “Design considerations” and “Additional Analyses” sections. Table I.15 and Figure I.5 show the result: the correction moves estimates modestly away from 0.5 without disturbing the ranking in which political attributes exert larger effects than demographic ones.

Table I.15: Measurement-error correction on directly feasible subsets

Attribute	k	Uncorrected μ [95% CI]	Corrected μ [95% CI]	Δ pp
Primary direct subset				
College Degree vs None	2	0.521 [0.503, 0.539]	0.523 [0.503, 0.544]	0.264
Conservative vs Liberal Issue	4	0.445 [0.287, 0.603]	0.443 [0.253, 0.633]	-0.176
Positions				
In-party vs Out-party	3	0.630 [0.580, 0.681]	0.651 [0.608, 0.693]	2.043
Majority vs Minority	2	0.526 [0.500, 0.551]	0.533 [0.500, 0.566]	0.735
Republican vs Democrat	3	0.479 [0.461, 0.497]	0.475 [0.458, 0.493]	-0.419
Women vs Men	2	0.627 [0.353, 0.901]	0.636 [0.336, 0.935]	0.914
Young vs Old	3	0.569 [0.482, 0.655]	0.575 [0.482, 0.669]	0.672
Expanded direct subset				
Christian vs None	2	0.497 [0.485, 0.509]	0.497 [0.483, 0.510]	-0.029
College Degree vs None	2	0.623 [0.569, 0.678]	0.637 [0.577, 0.696]	1.326
Conservative vs Liberal Issue	3	0.428 [0.388, 0.468]	0.420 [0.377, 0.463]	-0.779
Positions				
In-party vs Out-party	3	0.707 [0.661, 0.753]	0.730 [0.678, 0.781]	2.283
Majority vs Minority	2	0.513 [0.497, 0.529]	0.515 [0.497, 0.533]	0.156
No vs Any Political Experience	2	0.438 [0.401, 0.474]	0.431 [0.391, 0.471]	-0.663
Republican vs Democrat	3	0.447 [0.415, 0.478]	0.441 [0.406, 0.475]	-0.574
Women vs Men	3	0.508 [0.500, 0.516]	0.509 [0.500, 0.517]	0.068
Young vs Old	4	0.561 [0.537, 0.584]	0.567 [0.541, 0.592]	0.598

Estimates cover conjoints from all countries, not US studies only. Primary direct subset includes studies with at least 500 repeated-task pairs and at least 500 respondents contributing repeated identical tasks. Expanded direct subset includes studies with at least 100 repeated-task pairs and at least 100 respondents contributing repeated identical tasks. Corrected estimates use study-specific τ from repeated-task agreement. Corrected standard errors incorporate approximate uncertainty in τ via a delta-method propagation. Intervals are 95% confidence intervals for the meta-analytic mean.

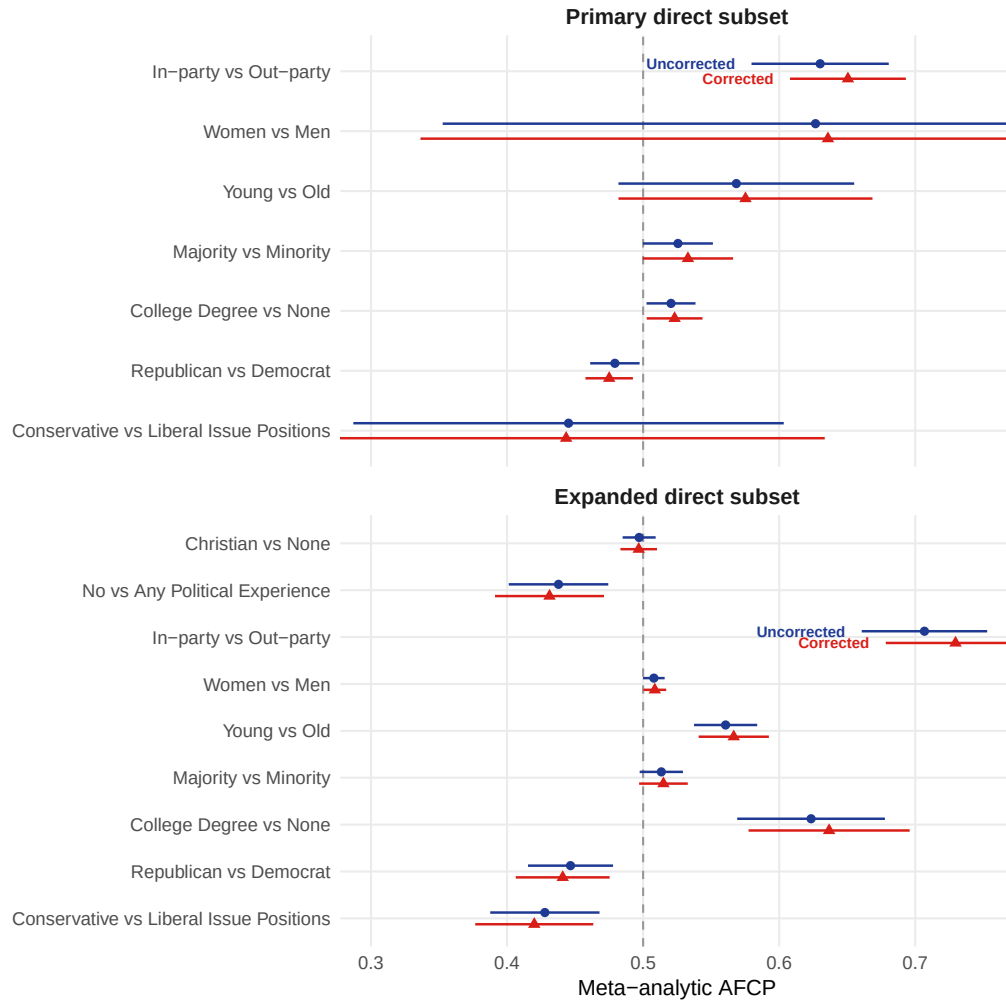


Figure I.5: Measurement-error correction on directly feasible subsets

Note: Estimates cover conjoints from all countries, not US studies only. The primary direct subset includes studies with at least 500 repeated-task pairs and at least 500 respondents contributing repeated identical tasks. The expanded direct subset includes studies with at least 100 repeated-task pairs and at least 100 respondents contributing repeated identical tasks. Corrected estimates use study-specific swapping-error parameters inferred from repeated-task agreement, and corrected confidence intervals propagate uncertainty in that parameter approximately via a delta-method calculation. The correction is a subset robustness check, not a full-corpus adjustment, because most studies do not include recoverable repeated identical tasks.

J AFCP intransitivity audit

Following Abramson et al. (2026), we audit whether the testable multi-level AFCP comparisons imply cyclic majority relations. This audit backs the main text’s “Additional Analyses” claim that the corpus is coherent rather than riddled with intransitivity. Table J.16 aggregates the check over every study-level triangle for which the corpus contains all three pairwise AFCP estimates. The pooled meta-analytic estimates are transitive in nine of the ten testable triangles. The exception, low versus middle versus upper income, cycles on pooled AFCPs of 0.506, 0.504, and 0.525 (upper over low), each with a confidence interval that includes 50%. At the study level, point-estimate cycles do appear, but almost all of them are weak and close to indifference; only two survive the stricter requirement that all three pairwise confidence intervals lie fully on one side of 50%. Table J.17 lists those two cycles and their pairwise win rates.

Table J.16: AFCP intransitivity audit across testable multi-level attribute triangles

Attribute	Triangle	Complete study-level triangles	Point cycles	Supported cycles	Pooled cycle
Age	Young / Middle / Old	125	36	0	No
Income	Low / Middle / Highest	18	1	0	No
Income	Low / Middle / Upper	19	6	0	Yes
Income	Low / Upper / Highest	18	1	0	No
Income	Middle / Upper / Highest	20	0	0	No
Party	Democrat / Independent / Republican	8	0	0	No
Policy	Liberal / Moderate / Conservative	73	7	3	No
Race	White / Asian / Hispanic	59	11	0	No
Race	White / Black / Hispanic	70	14	0	No
Religion	Christian / Muslim / None	21	0	0	No

Note: A study-level triangle is included when the corpus contains all three relevant pairwise AFCP estimates for a three-level comparison. A point cycle occurs when the pairwise majority relation implies $A \succ B$, $B \succ C$, and $C \succ A$ (or the reverse). A supported cycle additionally requires all three pairwise confidence intervals to lie fully on one side of 50%. Under that stricter rule, 3 supported cycles remain in the current corpus, in `mercea_santos_2025_study_3`, `sances_2018_study_1`, and `wuest_pontusson_2022_study_3`; their pairwise point estimates are shown in Table J.17.

Table J.17: Pairwise win rates for the supported study-level cycles

Study	Triangle ($A / B / C$)	$P(A \succ B)$	$P(B \succ C)$	$P(A \succ C)$
mercea santos 2025 study 3	Liberal / Moderate / Conservative	0.459	0.457	0.586
sances 2018 study 1	Liberal / Moderate / Conservative	0.205	0.424	0.675
wuest pontusson 2022 study 3	Liberal / Moderate / Conservative	0.405	0.440	0.568

Note: Estimates cover conjoints from all countries, not US studies only. Each row is a study-level triangle whose three pairwise AFCPs form a cycle with all three confidence intervals off 50%. A , B , and C are the three attribute levels in the order listed. A cycle means $P(A \succ B) > 0.5$, $P(B \succ C) > 0.5$, and $P(A \succ C) < 0.5$ (or all three reversed): the pairwise majorities do not agree on a single ordering.

K Publication bias

The meta-reanalysis design offers unusual protection against file-drawer selection. As documented in the main text, for most of the matchup estimates in our corpus we could find no directly comparable published estimate: we recompute each win rate from a study’s raw data whether or not, and in whatever form, its authors presented it. The usual mechanism of publication bias, in which results that fail to reach significance are less likely to appear in print, therefore has little purchase at the level of the individual matchup, because the estimate’s existence in our corpus does not depend on its significance or its sign.

We verify this directly with a three-parameter selection model (Vevea and Hedges, 1995), fit separately to each matchup over every study that randomized the relevant levels, pooling across countries rather than splitting by them. Two matchups are US-specific in practice rather than by construction: all 34 Republican-versus-Democrat estimates and 43 of the 45 white-versus-Black estimates come from US studies, because almost no study elsewhere randomizes those levels. The three parameters are a pooled win rate, a between-study heterogeneity SD $\hat{\tau}$, and a selection weight $\hat{\lambda}$. We place the selection step at the two-sided $p = 0.05$ boundary against the 0.5 null, so that a “significant” study is one whose interval excludes indifference in either direction. The weight $\hat{\lambda}$ is then the probability that a non-significant estimate enters the corpus relative to a significant one: $\hat{\lambda} = 1$ signals no selection, $\hat{\lambda} < 1$ the file-drawer pattern in which significant results crowd out null ones, and $\hat{\lambda} > 1$ the reverse. Because significance is defined two-sided, $\hat{\lambda}$ does not depend on the sign or the predominant direction of a matchup’s effect.

Table K.18 reports the results, and they are reassuring. The selection-adjusted win rates track the unadjusted estimates throughout: the headline comparisons scarcely move (in-party $72.5 \rightarrow 72.5$, women $53.0 \rightarrow 53.5$, no political experience $43.2 \rightarrow 42.6$), and the largest adjustment anywhere is under a point and a half. Not one significant matchup shows the file-drawer signature of $\hat{\lambda} < 1$. The three matchups whose selection tests reach significance (women versus men, majority versus minority, and white versus Black) point the other way, toward an over-representation of *non*-significant estimates ($\hat{\lambda} = 1.7, 1.6$, and 4.4). The selection those weights describe is substantial; its consequence for the pooled estimate is not, because a matchup’s studies are numerous and its heterogeneity modest, which is why the adjusted means above barely move. As it happens, that is exactly what the meta-reanalysis design should produce: because every matchup is recomputed from the raw data, null results survive into our corpus that a significance-filtered literature would have discarded. In short, our estimates are not materially distorted by publication selection.

Table K.18: Publication-bias check: two-sided selection model on the study-level AFCPs, pooled across countries

Matchup	k	Pooled AFCP (%)			$\hat{\tau}$ (pp)	$\hat{\lambda}$	Selection p
		Unadjusted	Adjusted	95% CI			
In-party vs Out-party	79	72.5	72.5	[70.6, 74.4]	8.1	0.83	0.870
Republican vs Democrat	49	46.8	46.7	[45.2, 48.1]	4.2	1.12	0.762
Conservative vs Liberal Issue Positions	95	42.9	43.4	[41.4, 45.4]	8.8	0.71	0.233
Women vs Men	218	53.0	53.5	[52.8, 54.2]	4.0	1.71	0.002
Young vs Old	114	57.5	57.6	[56.4, 58.8]	4.8	1.08	0.791
Majority vs Minority	106	52.6	52.9	[51.4, 54.4]	6.9	1.64	0.039
White vs Black	61	48.4	46.9	[44.8, 49.0]	4.7	4.37	<0.001
College Degree vs None	55	56.6	57.2	[55.4, 59.0]	5.5	1.76	0.172
No vs Any Political Experience	120	43.2	42.6	[41.3, 43.9]	5.8	1.56	0.104

Note: Estimates cover conjoints from all countries, not US studies only. Three-parameter step selection model (Vevea and Hedges, 1995), estimated separately for each matchup on the study-level AFCPs, pooling every study that randomized the relevant levels rather than splitting by country. Two matchups are US-specific in practice rather than by construction: every Republican-versus-Democrat estimate and all but two of the white-versus-Black estimates come from US studies. The step is placed at the two-sided $p = 0.05$ significance boundary against the 0.5 null, so “significance” means a study’s 95% interval excludes 0.5 in either direction; the model does not depend on the sign or predominant direction of a matchup’s effect. The adjusted mean re-estimates the pooled AFCP allowing significant and non-significant estimates to enter the sample at different rates. k is the number of contributing study-level estimates, which exceeds the number of studies wherever a study randomized whether policy information is shown and so contributes one estimate per condition. $\hat{\tau}$ is the between-study heterogeneity SD (pp). $\hat{\lambda}$ is the estimated selection weight: the publication probability of a non-significant estimate relative to a significant one, with $\hat{\lambda} = 1$ denoting no selection, $\hat{\lambda} < 1$ under-representation of non-significant estimates (file-drawer bias), and $\hat{\lambda} > 1$ over-representation. Selection p is the likelihood-ratio test of the null of no selection ($\lambda = 1$).

L Party–policy tradeoffs by respondent ideology

The main text’s party–policy figure classifies a candidate’s policy positions as congruent when they match the respondent’s party, because most studies do not measure respondents’ own issue positions. Many do measure ideology, however. Twenty-one of the 27 US studies in that figure include a liberal–conservative self-placement item (the exceptions are both samples in Bansak et al. (2021), Ferrari and Smith (2025), Loewen and Rheault (2021), and both Peterson (2017) samples); 19 of the 21 contribute to the key party-versus-policy comparison in the lower-left panel. We split Democratic and Republican respondents into liberals, moderates, and conservatives using each study’s own response labels. Every one of these scales has an odd number of points with a labeled midpoint, so the classification requires no cutpoint choices: “slightly liberal” counts as liberal, the midpoint as moderate, and “slightly conservative” as conservative. We then re-estimate that figure’s six comparisons within each of the six party-by-ideology groups and pool them across studies with the same random-effects model. Ideology is measured, not randomized, so differences across these groups describe who makes which tradeoff; they do not estimate the effect of ideology itself.

The key comparison (lower-left panel) separates two kinds of partisans. Ideologically consistent partisans lean toward the out-partisan who holds their party’s typical positions: liberal Democrats choose the co-partisan with conservative positions 38.2% of the time (95% CI: 31.7–44.7%, $k = 18$), and conservative Republicans choose the co-partisan with liberal positions 47.7% of the time (95% CI: 39.9–55.4%, $k = 19$). Cross-pressured partisans reverse the pattern: conservative Democrats choose the co-partisan 63.6% of the time (95% CI: 58.9–68.3%, $k = 14$) and liberal Republicans 61.4% of the time (95% CI: 52.2–70.6%, $k = 12$). Moderates fall between the two (Democrats 45.6%, 95% CI: 40.2–51.0%; Republicans 51.6%, 95% CI: 44.1–59.1%). For cross-pressured partisans the co-partisan’s “mismatched” positions are the ones closer to their own ideology, so both groups choose the candidate whose positions match their own ideology. The other five panels show the same ordering. The co-partisan advantage when policy is held constant is largest for liberal Democrats and conservative Republicans, and the preference for party-typical positions when party is held constant is weakest, and sometimes reversed, among cross-pressured partisans.

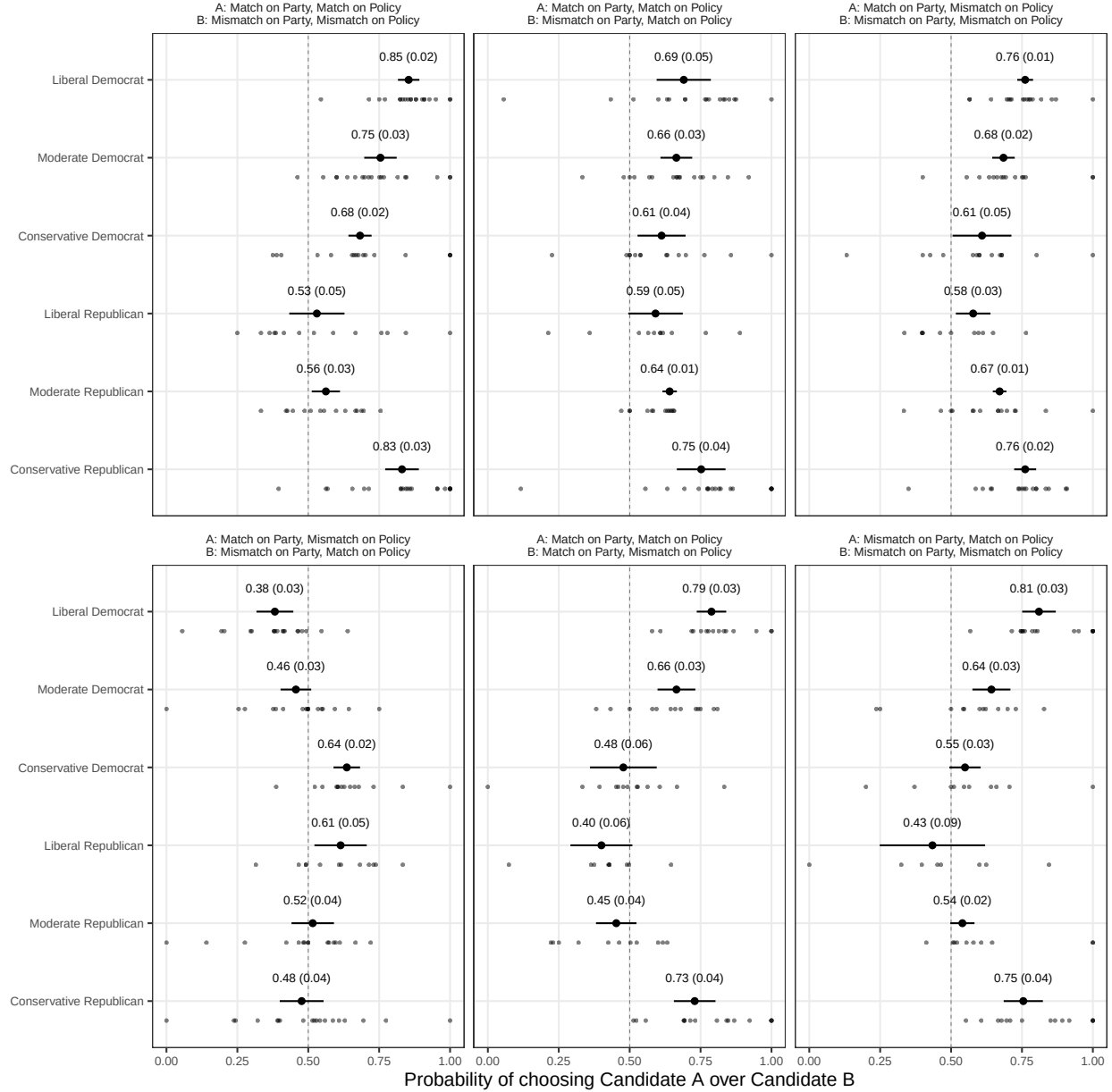


Figure L.6: Party-policy tradeoffs by respondent partisanship and ideology

Note: Same six comparisons as the main text's party-policy figure, with Democratic and Republican respondents each split into liberals, moderates, and conservatives by their answer to the study's ideology self-placement item. Policy congruence is still defined by the respondent's party, not their ideology. Restricted to the 21 US studies in that figure that measure ideology; k varies by panel and group because some studies have too few respondents in a cell. Small points below each row label are individual study-level AFPC estimates; large points are random-effects meta-analytic means with 95% confidence intervals. The dashed vertical line marks 0.5 (indifference between candidates A and B).

M Party–policy tradeoffs by respondents’ own issue positions

The ideology split in Appendix L measures congruence against a general liberal–conservative label rather than the specific issues on the candidate profiles. Nine of the US studies in the main text’s party–policy figure also asked respondents for their own positions on the issues the candidates took stands on: Costa (2021), Funck and McCabe (2022), Graham and Svolic (2020), Leeper and Robison (2020), Miller and Ziegler (2024), Mummolo, Peterson, and Westwood (2021), both Peterson (2017) samples, and Sevi et al. (2025). For these studies we code a candidate’s position on each displayed issue as matching the respondent’s when the two sit on the same side of it; a midpoint answer, whether the candidate’s or the respondent’s, matches neither way. A candidate matches on policy when every displayed issue on which the respondent took a side agrees with the respondent, and mismatches when every one disagrees; a candidate with mixed agreement drops out of the comparisons. We then re-estimate the six comparisons of the main-text figure with this definition of policy match, separately for Democrats and Republicans.

The key comparison (lower-left panel) pits a co-partisan who disagrees with the respondent against an out-partisan who agrees. Democrats choose the co-partisan 41.7% of the time (95% CI: 38.2–45.2%, $k = 9$) and Republicans 43.0% of the time (95% CI: 40.0–46.0%, $k = 9$). In the main-text figure, where congruence is defined by the respondent’s party and the estimates draw on many more studies, the Republican interval includes 0.5; here it does not. Measured against their own positions, partisans of both parties prefer the out-partisan who agrees with them on the issues. The other five panels look like the main text: holding policy match constant, the co-partisan wins 70.2–73.9% of the time, and holding party constant, the candidate who agrees with the respondent wins 73.2–77.5% of the time. The two within-party comparisons rest on the four studies that supply them.

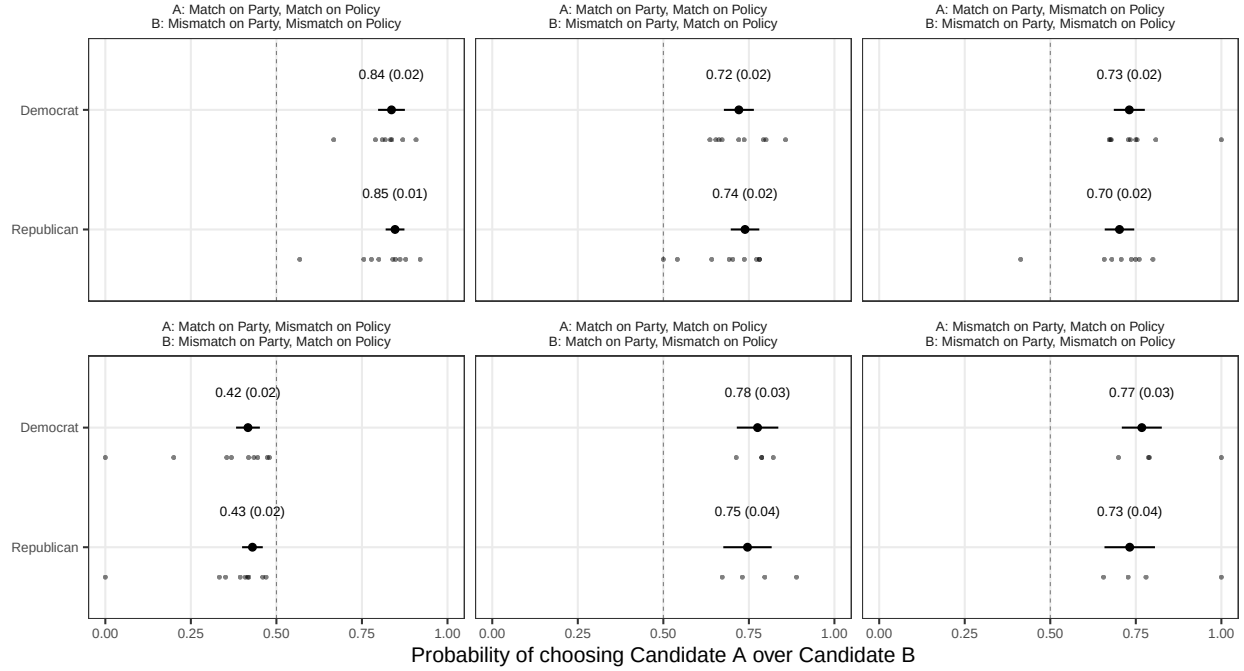


Figure M.7: Party-policy tradeoffs with policy congruence defined by the respondent's own issue positions

Note: Same six comparisons as the main text's party-policy figure, except that a candidate matches on policy when every displayed issue on which the respondent took a side agrees with the respondent's own stated position, and mismatches when every one disagrees. Restricted to the nine US studies in that figure that asked respondents about the candidates' issues; k is 9 in most panels (8 in two cells where one study contributes no estimate) and 4 in the two within-party comparisons. Small points below each row label are individual study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. The dashed vertical line marks 0.5 (indifference between candidates A and B).

N Party-demographic forest plots

This section provides the study-by-study evidence behind the summary of candidate choice tradeoffs in the main text’s “Cross-pressured Respondents” section, whose final figure pools these results into a single grid. Each figure here is a forest plot for one cell of that analysis: it fixes a respondent subgroup (defined by party and by gender or race) and a pair of candidate types that match or mismatch the respondent on party and on the demographic attribute, and shows the study-level AFCP of choosing candidate A alongside the random-effects meta-analytic mean. Reading down a figure shows how consistent the tradeoff is across studies. The six figures for each attribute walk through every combination of party-match and demographic-match, in the order both/neither, both/party, both/demographics, party/neither, party/demographics, demographics/neither. The fifth in each series, party versus demographics (Figure N.12 for gender and Figure N.18 for race), is the core tradeoff in which one candidate matches only on party and the other only on the demographic. Throughout, gold points come from studies that show policy information and blue points from those that do not.

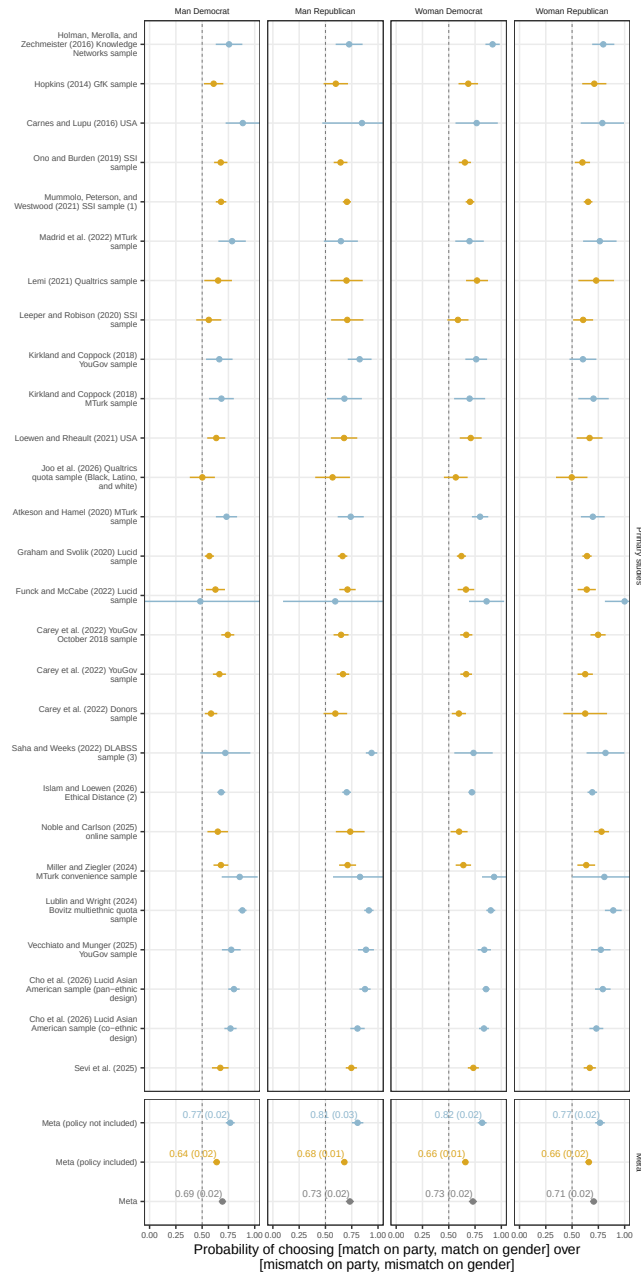
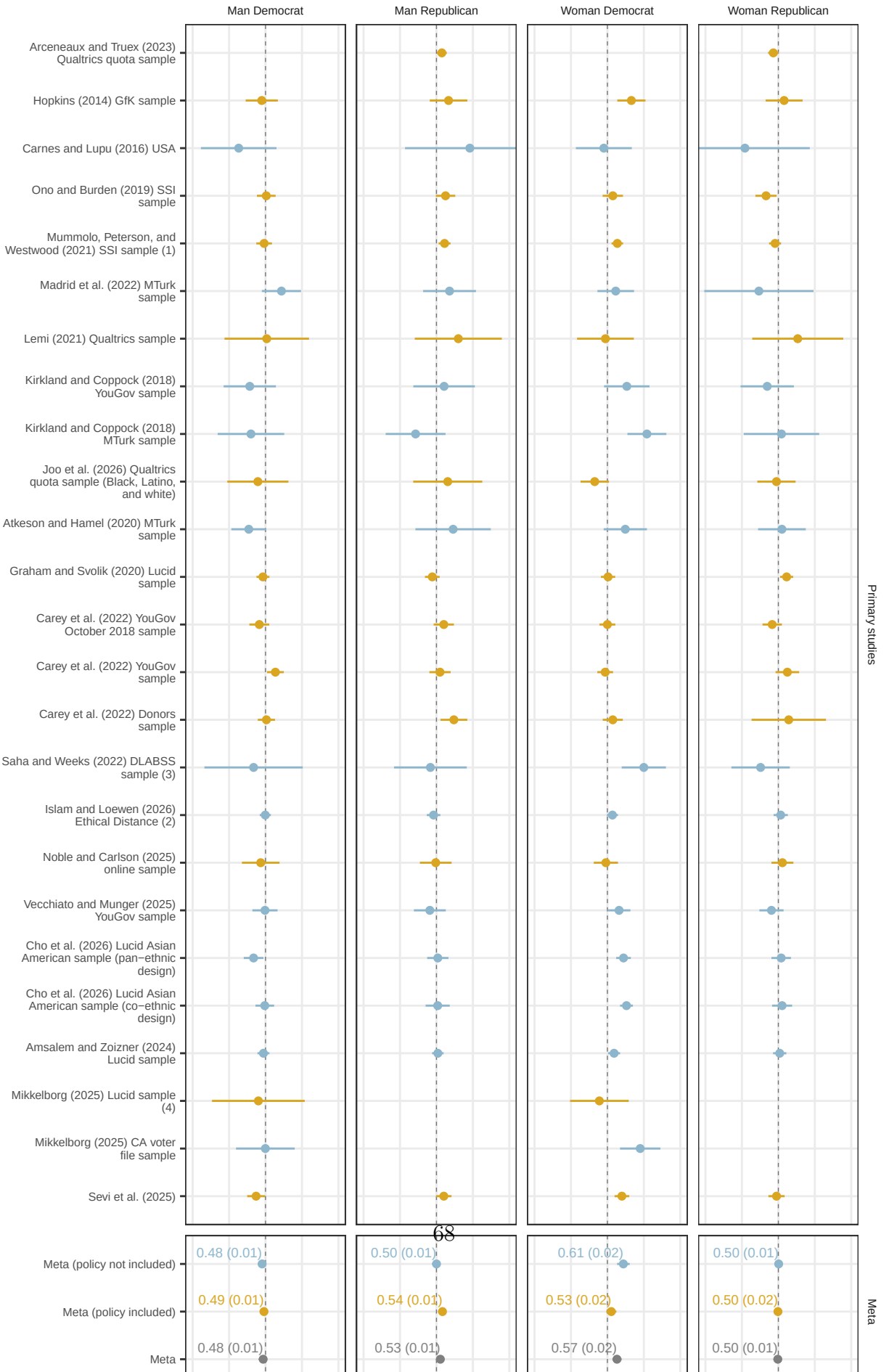


Figure N.8: Comparison of both versus neither (gender)

Note: Estimates are restricted to US studies. Candidate A matches the respondent on both party and gender; Candidate B matches on neither. Rows are respondent subgroups defined by party and gender. Small points are study-level AFCEP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.



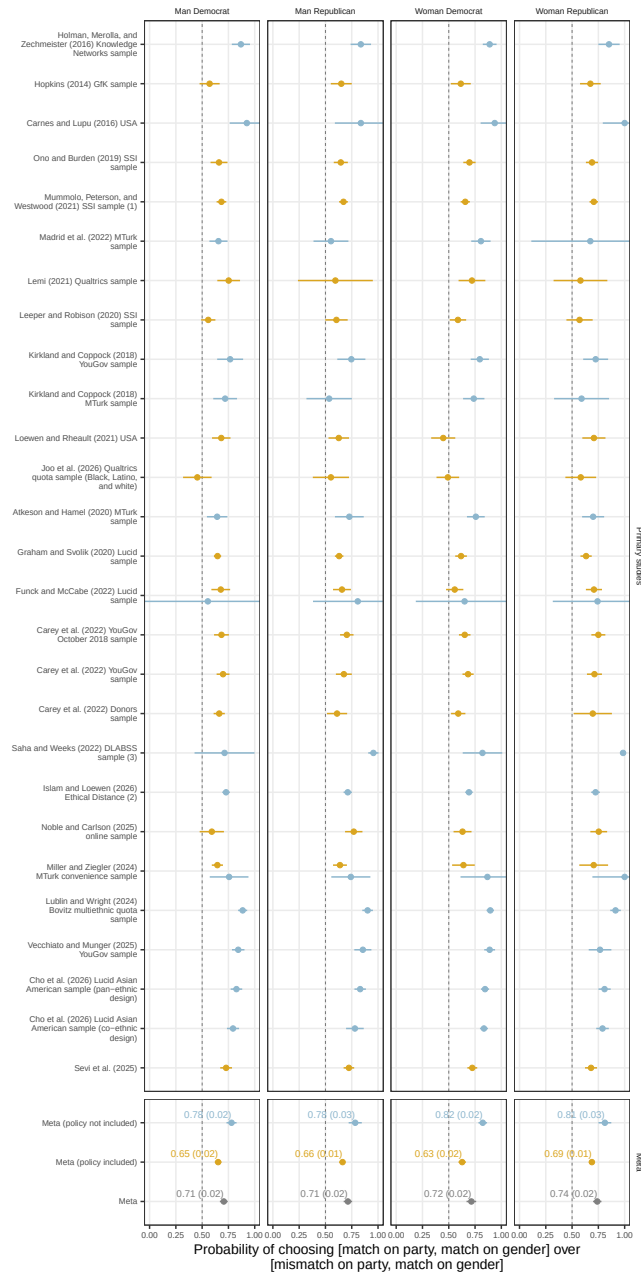


Figure N.10: Comparison of both versus demographics (gender)

Note: Estimates are restricted to US studies. Candidate A matches on both party and gender; Candidate B matches only on gender. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

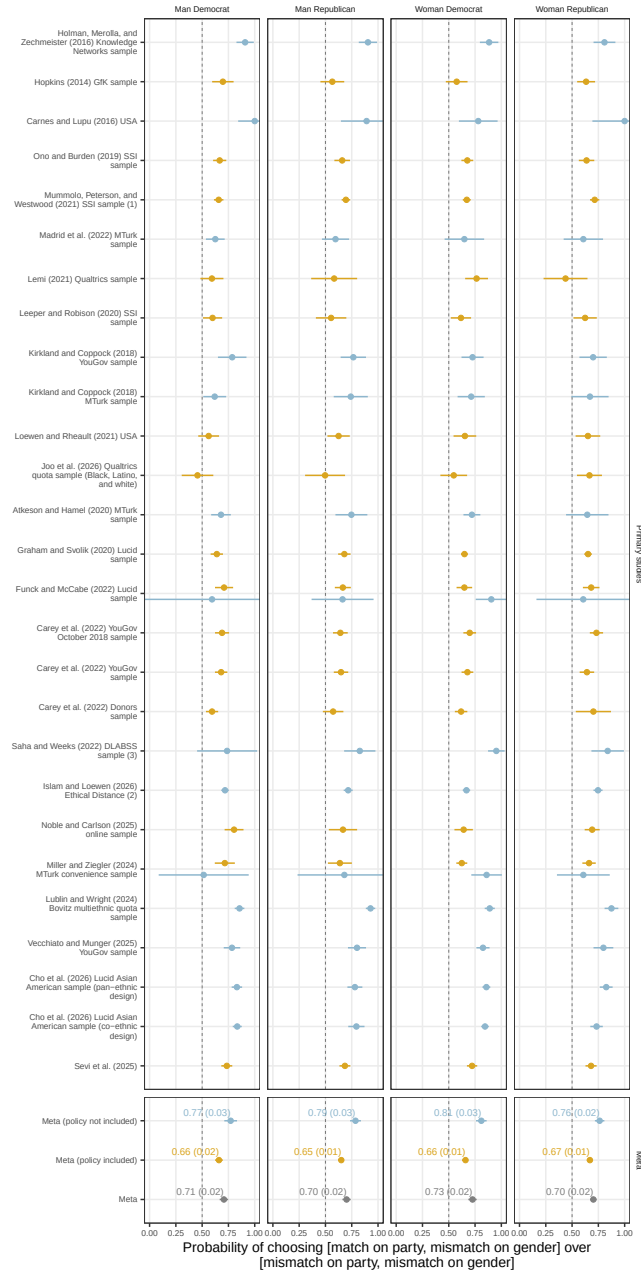


Figure N.11: Comparison of party versus neither (gender)

Note: Estimates are restricted to US studies. Candidate A matches only on party; Candidate B matches on neither party nor gender. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

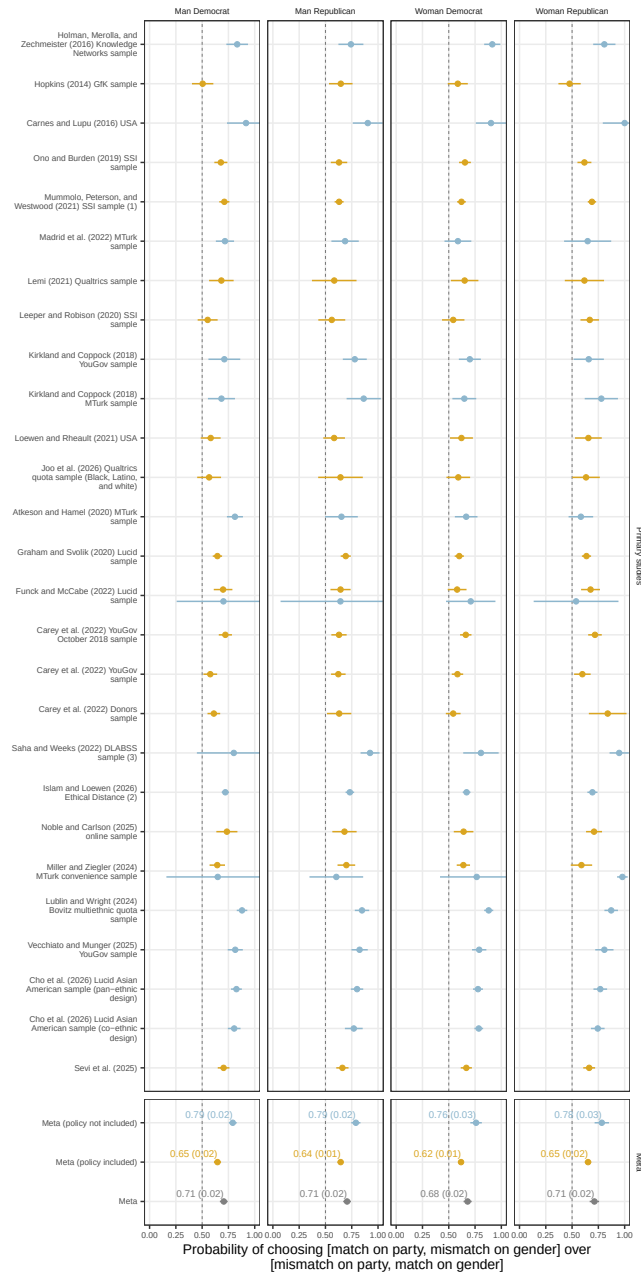
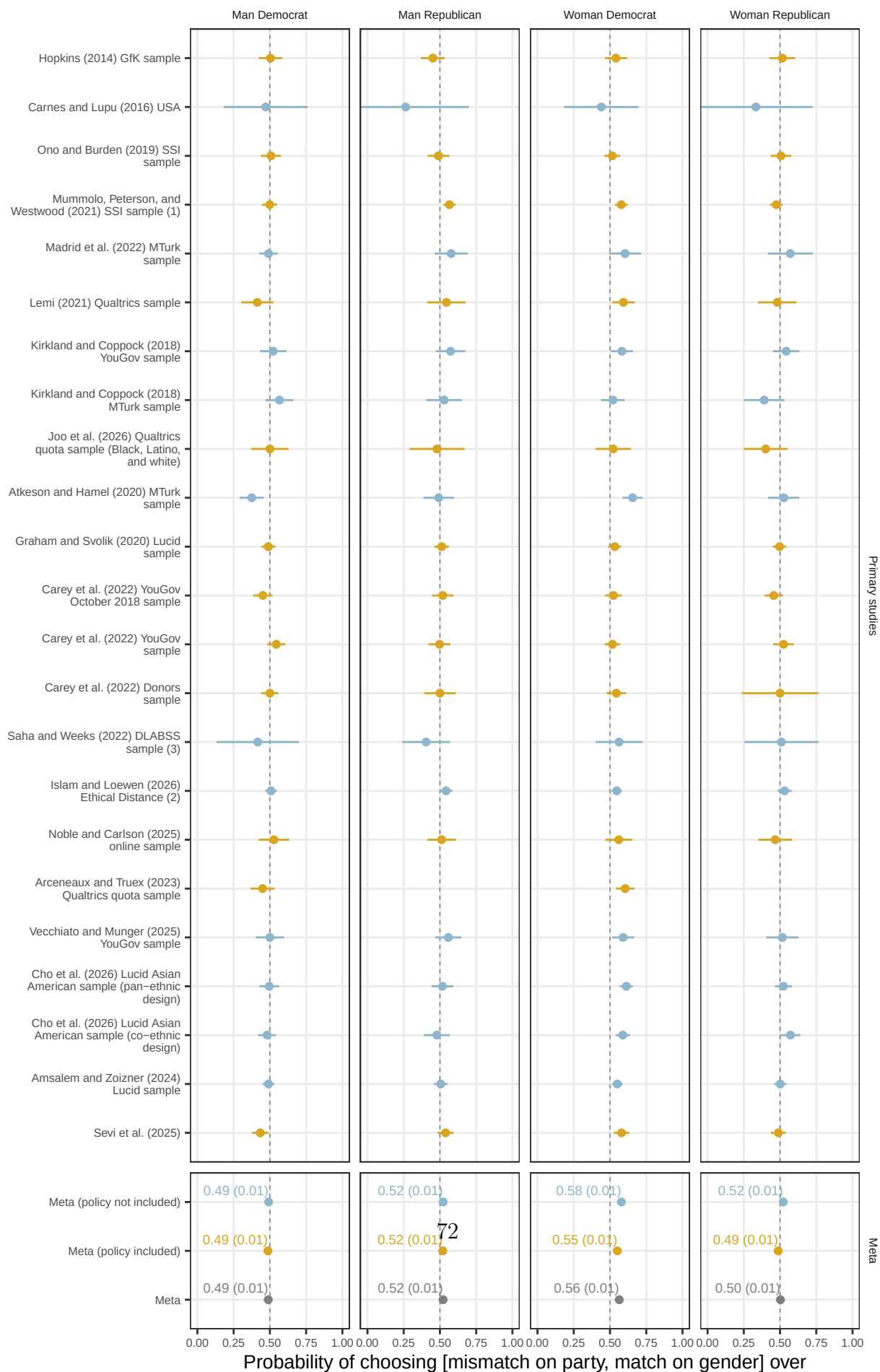


Figure N.12: Comparison of party versus demographics (gender)

Note: Estimates are restricted to US studies. Candidate A matches the respondent on party but not gender; Candidate B matches on gender but not party. This comparison is the core tradeoff. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.



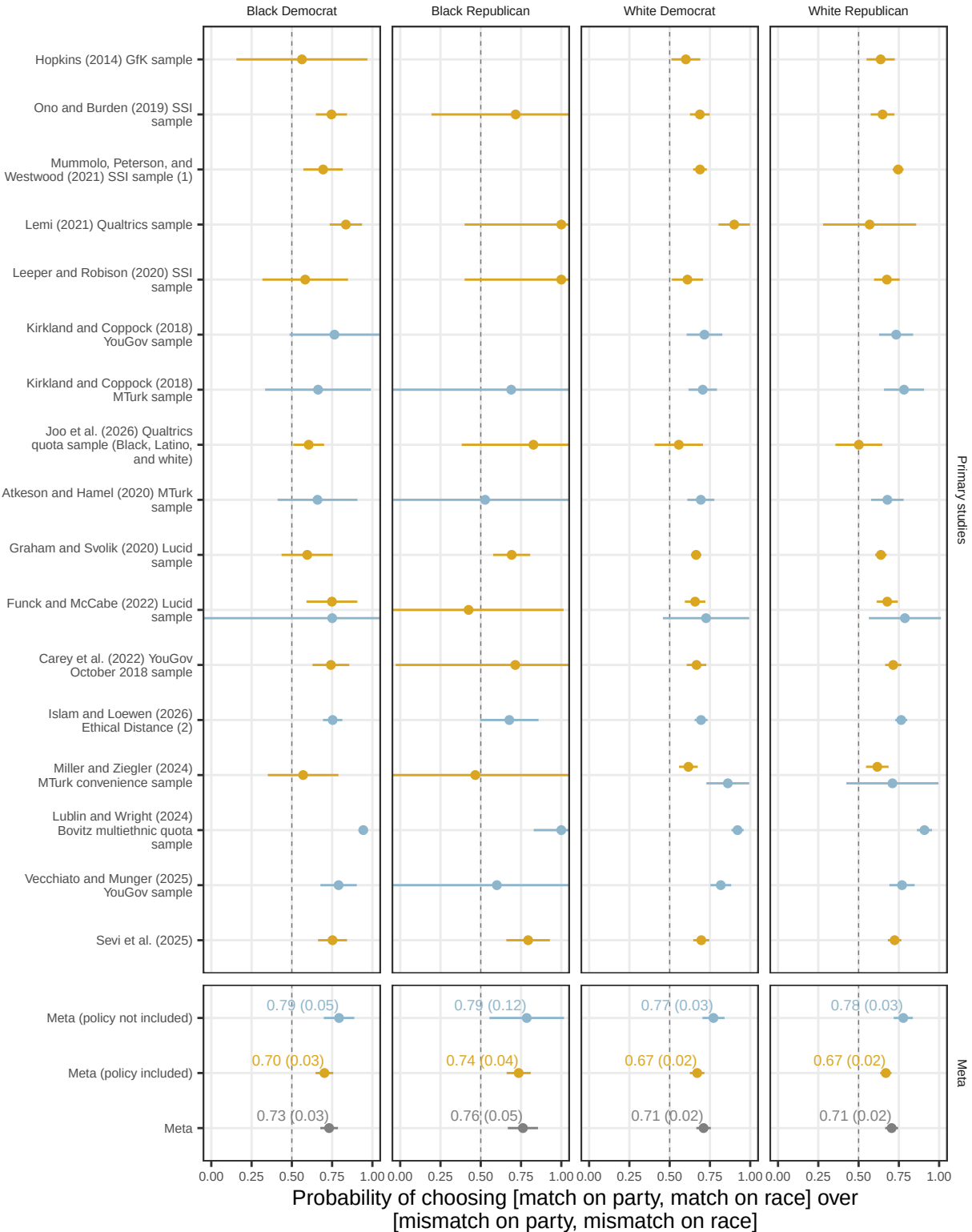


Figure N.14: Comparison of both versus neither (race)

Note: Estimates are restricted to US studies. Candidate A matches the respondent on both party and race; Candidate B matches on neither. Rows are respondent subgroups defined by party and racial group. Small points are study-level AFCEP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

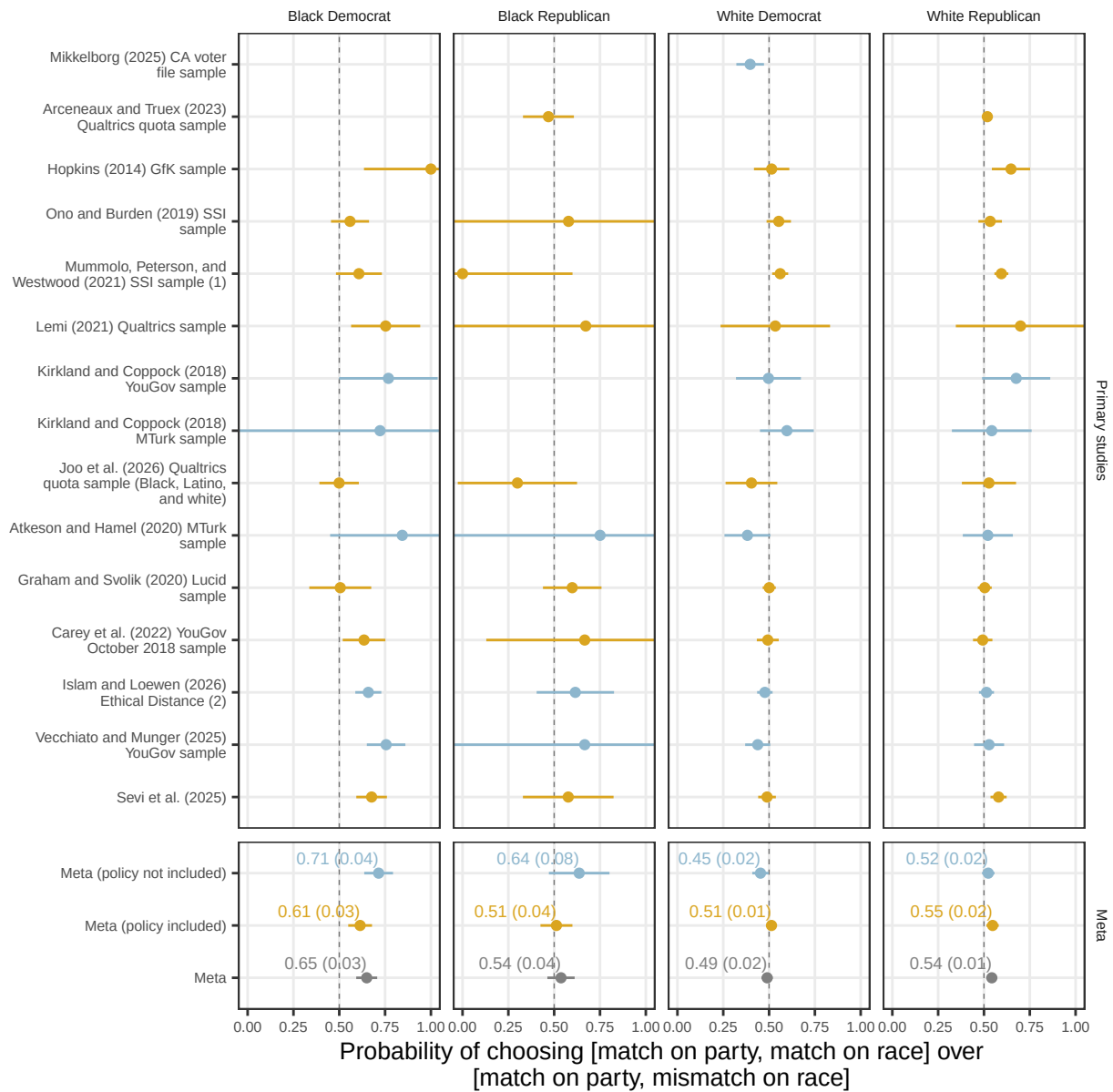


Figure N.15: Comparison of both versus party (race)

Note: Estimates are restricted to US studies. Candidate A matches on both party and race; Candidate B matches only on party. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

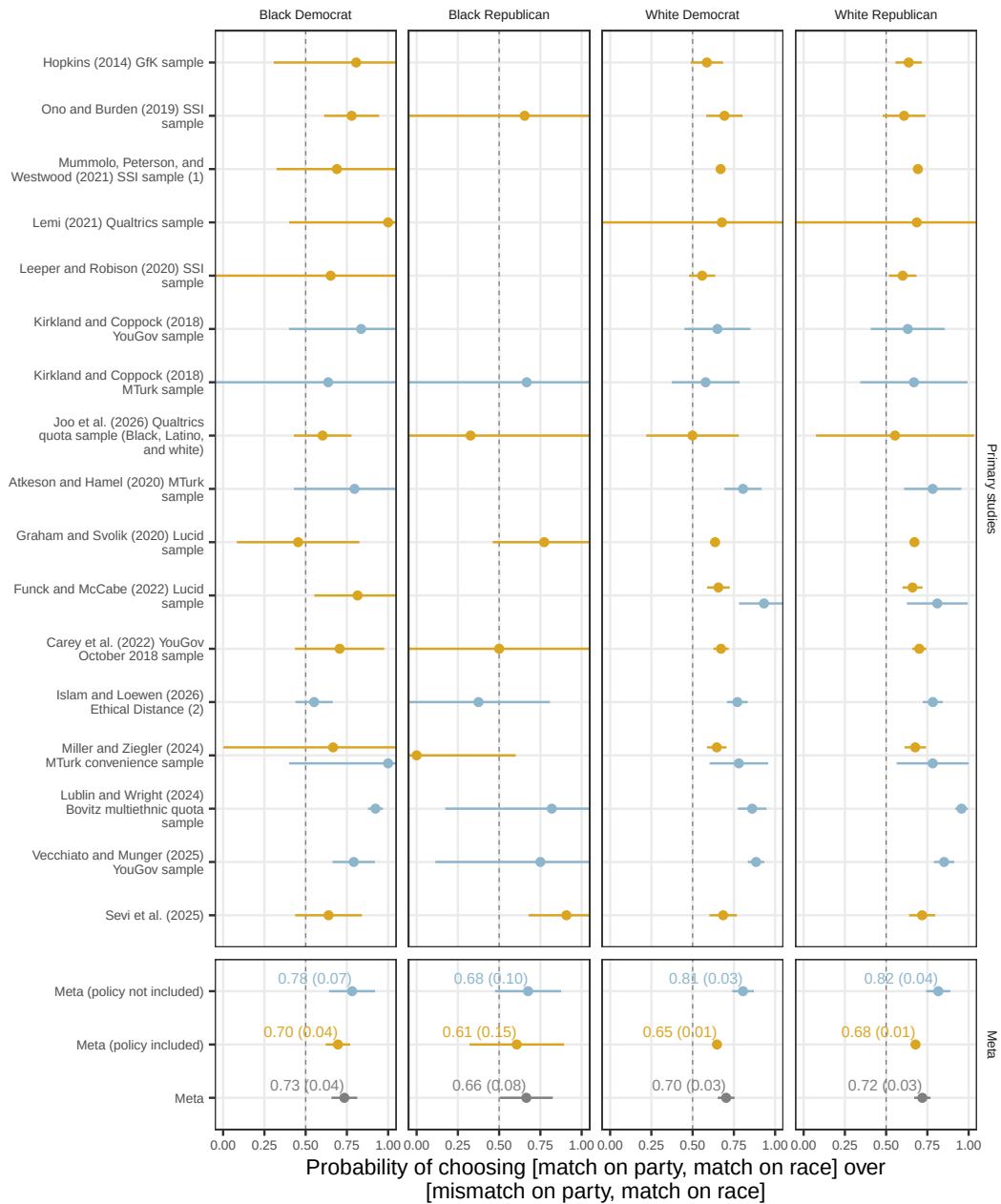


Figure N.16: Comparison of both versus demographics (race)

Note: Estimates are restricted to US studies. Candidate A matches on both party and race; Candidate B matches only on race. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

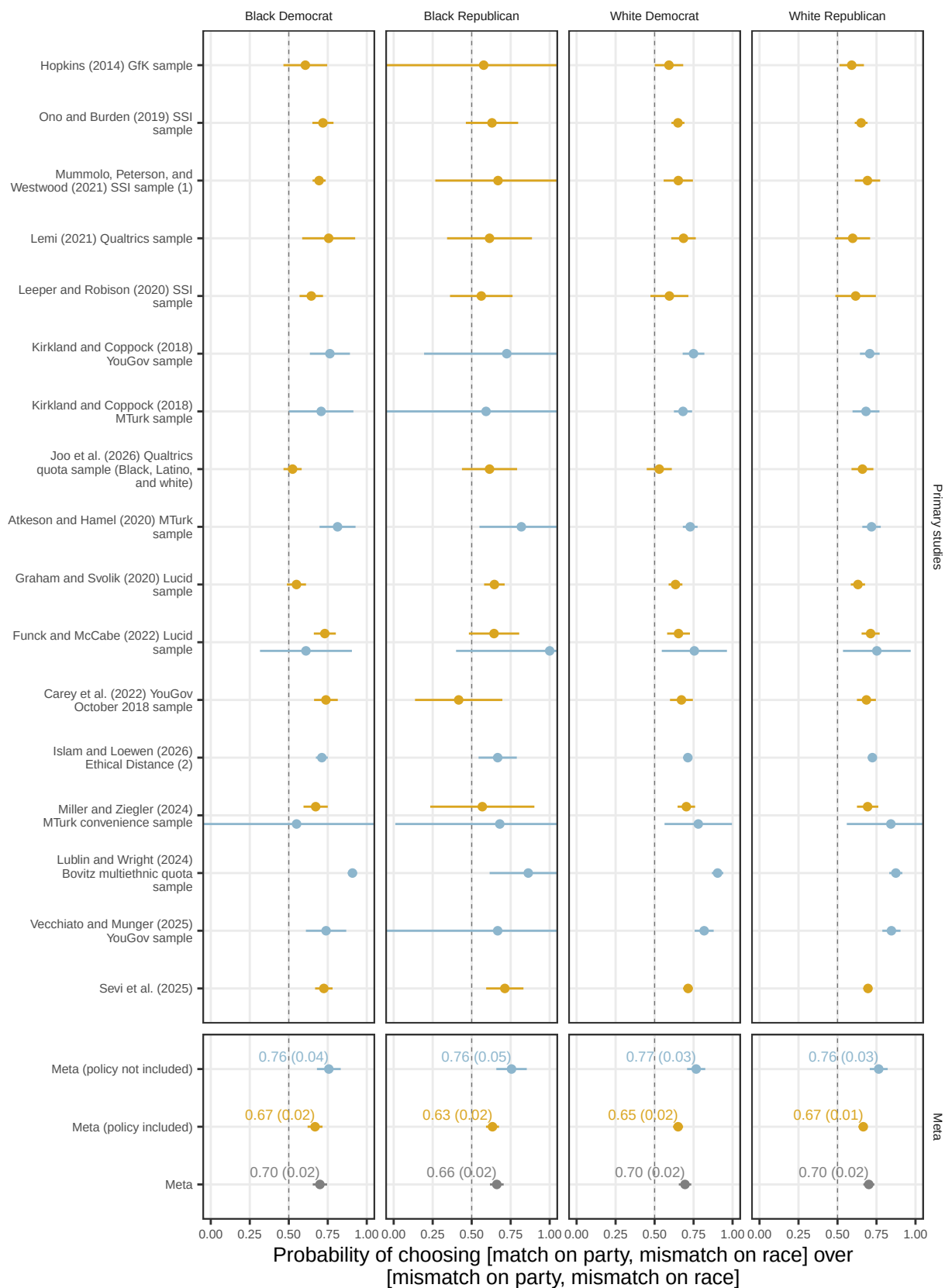


Figure N.17: Comparison of party versus neither (race)

Note: Estimates are restricted to US studies. Candidate A matches only on party; Candidate B matches on neither party nor race. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

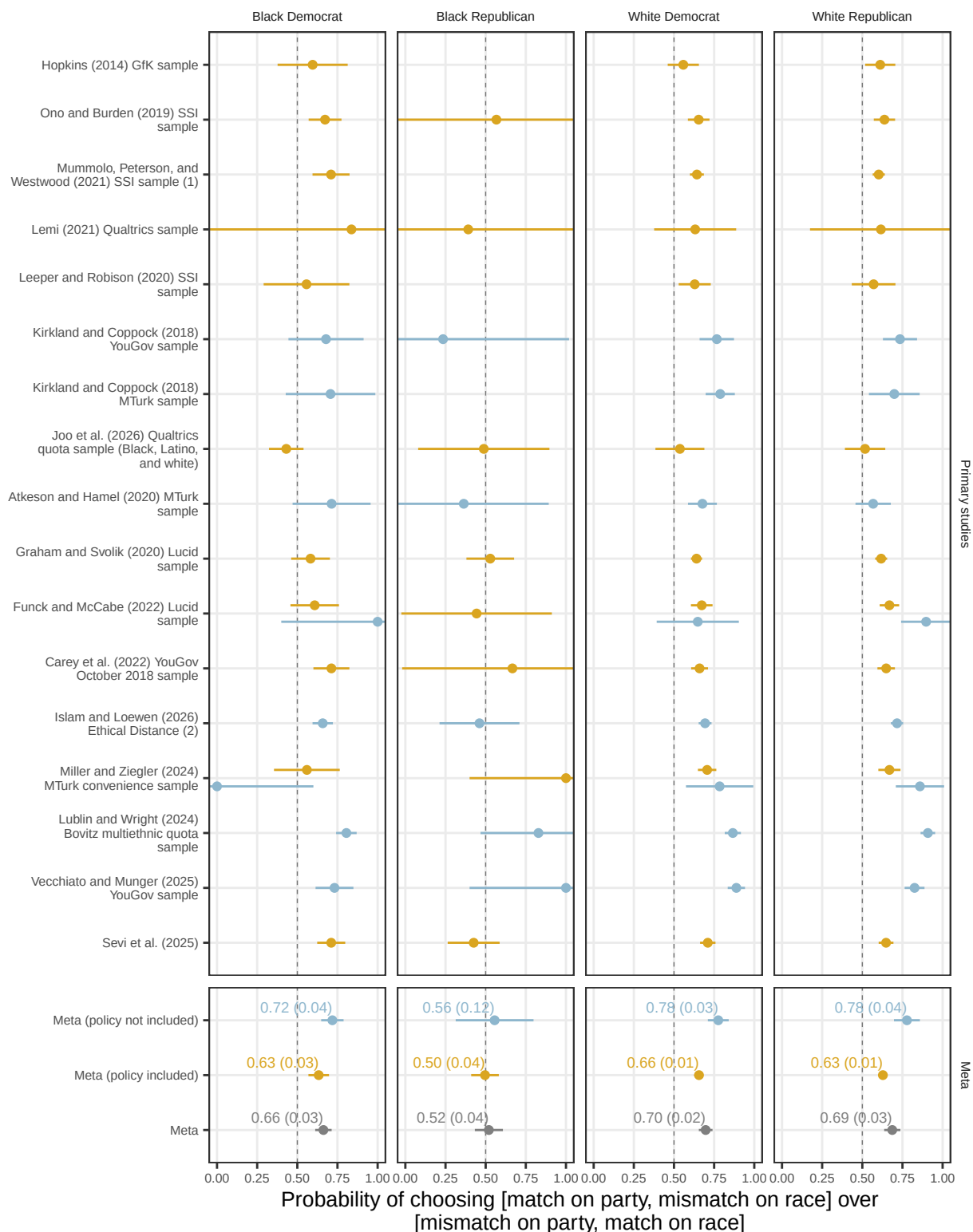


Figure N.18: Comparison of party versus demographics (race)

Note: Estimates are restricted to US studies. Candidate A matches the respondent on party but not race; Candidate B matches on race but not party. This comparison is the core tradeoff. Rows are respondent subgroups. Small points are study-level AFCP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

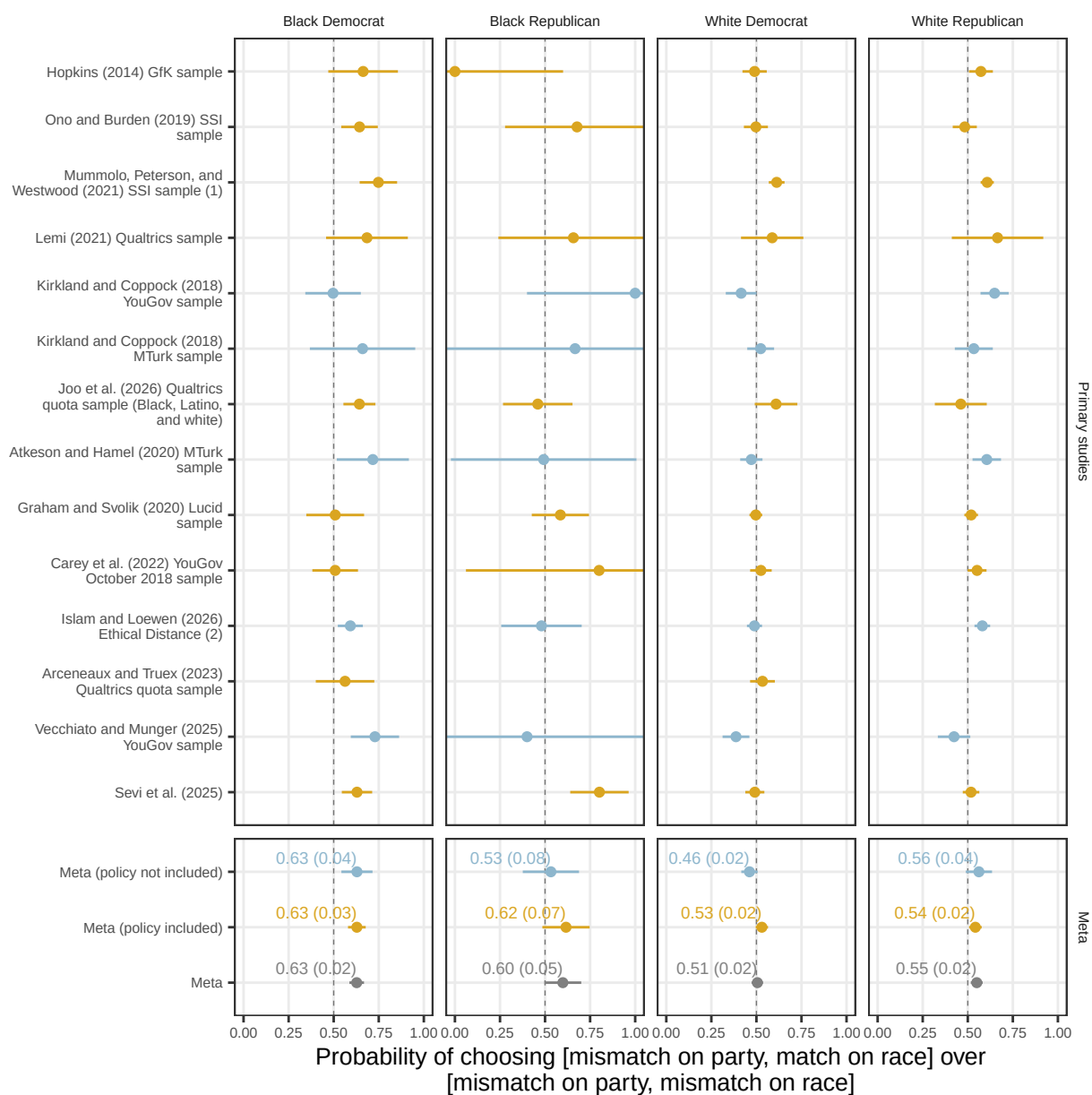


Figure N.19: Comparison of demographics versus neither (race)

Note: Estimates are restricted to US studies. Candidate A matches only on race; Candidate B matches on neither party nor race. Rows are respondent subgroups. Small points are study-level ACP estimates; large points are random-effects meta-analytic means with 95% confidence intervals. Gold: studies with policy information; blue: without.

O Per-attribute forest plots

Each figure below shows the distribution of study-level win-rate estimates for a single matchup, together with the random-effects meta-analytic mean and its 95% confidence interval, drawn in black at the top of each panel. The figures are grouped by the paper’s three attribute families: demographics/identity, experience, and politics. Most US figures split estimates by respondent partisanship (Democrats and Republicans shown separately). The race subsection below also includes a dedicated respondent-race split for white-versus-Black candidate matchups. Together these forests give the study-by-study evidence behind the average effects reported in the main text’s Results, and they reproduce the prior single-attribute meta-analyses cited there.

Demographics and identity

The following figure aggregates the international evidence to the country level for the two most commonly discussed demographic matchups. It is designed to show where the pooled averages are summarizing broadly similar country patterns and where they are masking meaningful cross-national spread.

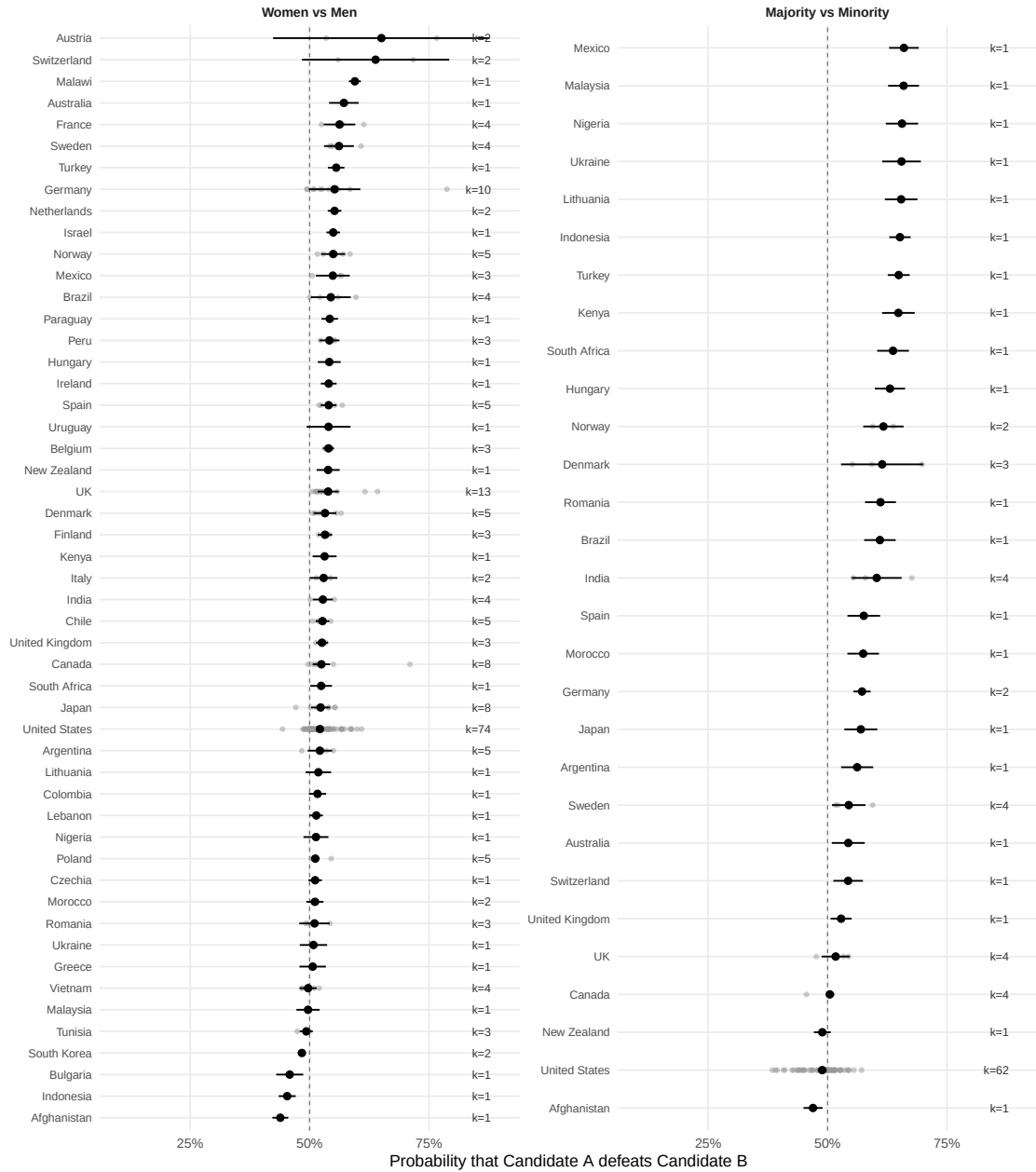


Figure O.20: Country-level breakdowns for gender and race matchups

Note: Estimates cover conjoints from all countries, not US studies only. The left panel shows country-level estimates for women versus men; the right panel shows majority versus minority race. Small gray points are study-level AFCP estimates. Black points and intervals are country-level means with 95% confidence intervals, using a random-effects aggregation when multiple studies are available within country. The dashed line marks 0.5. Labels at right give the number of contributing study-level estimates in each country.

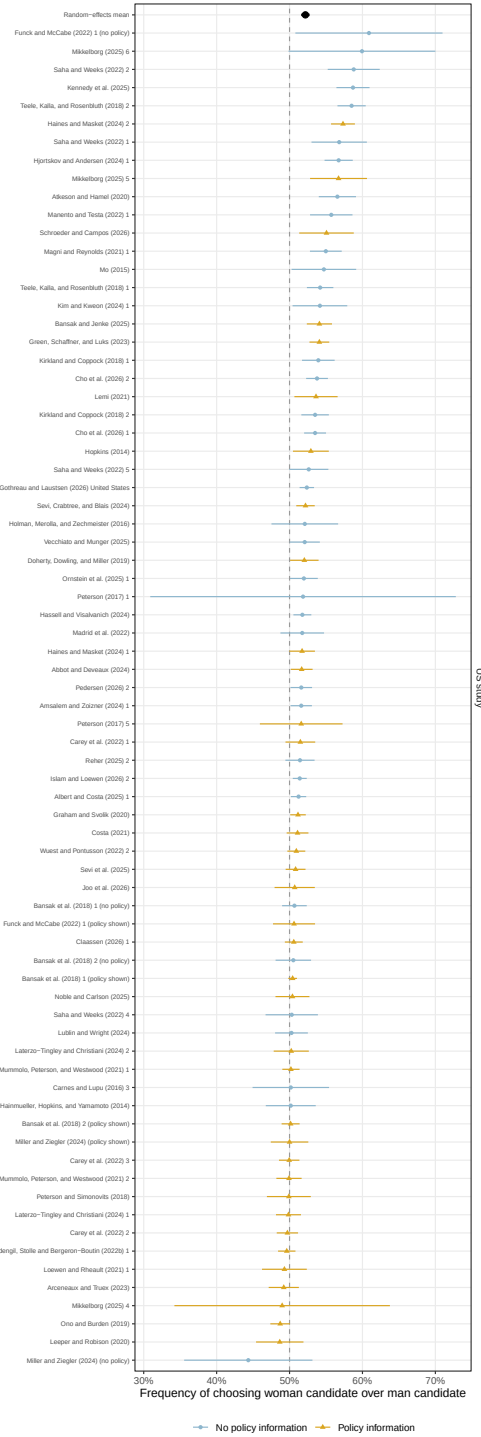


Figure O.21: Women versus Men win rates (US studies)

Note: Each row is one US study estimate with its 95% confidence interval. Points are colored and shaped by whether the study shows policy information. The black point at the top is $\hat{\mu}$ with its 95% confidence interval from a random-effects policy model over the US studies. The dashed line marks 0.5. The 218 study-level estimates for this matchup, from 212 studies, are shown across three figures for legibility: US studies here, and the non-US studies split alphabetically across the next two figures.

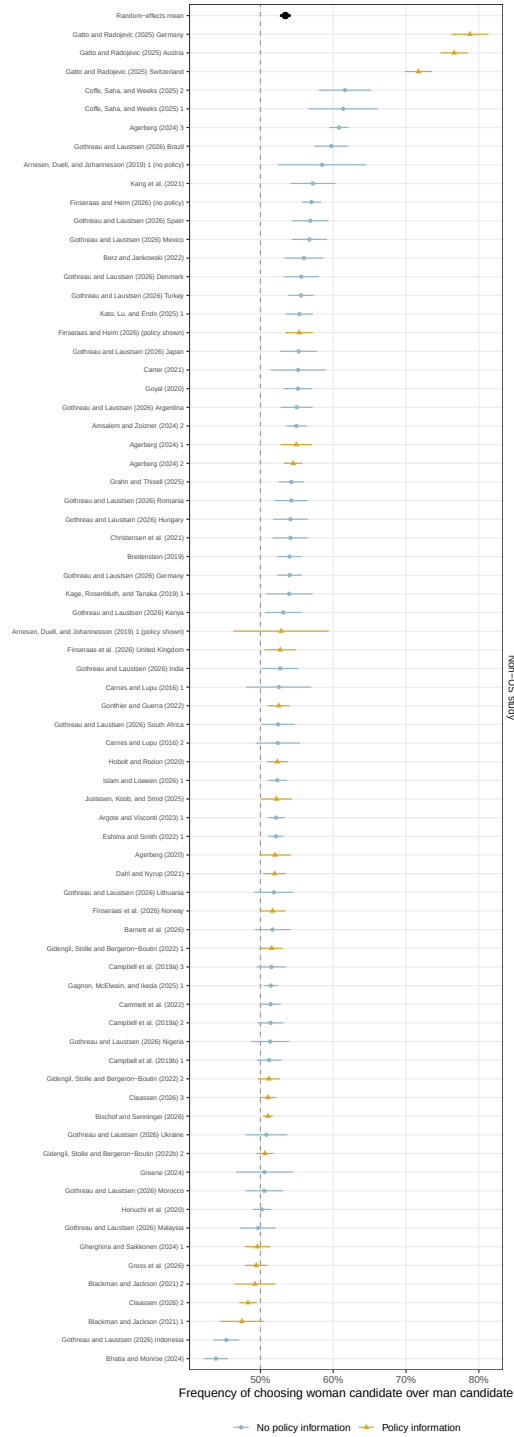


Figure O.22: Women versus Men win rates (non-US studies, part 1 of 2)
Note: Each row is one non-US study estimate with its 95% confidence interval, for the first alphabetical half of the non-US studies. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5.

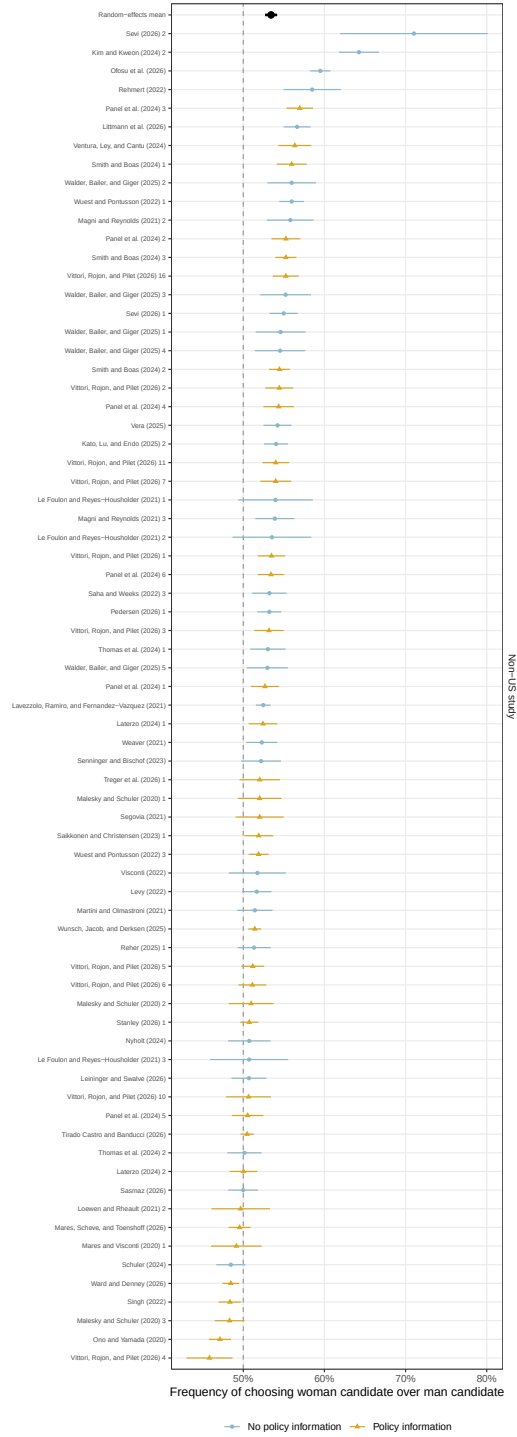


Figure O.23: Women versus Men win rates (non-US studies, part 2 of 2)
Note: Each row is one non-US study estimate with its 95% confidence interval, for the second alphabetical half of the non-US studies. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5.

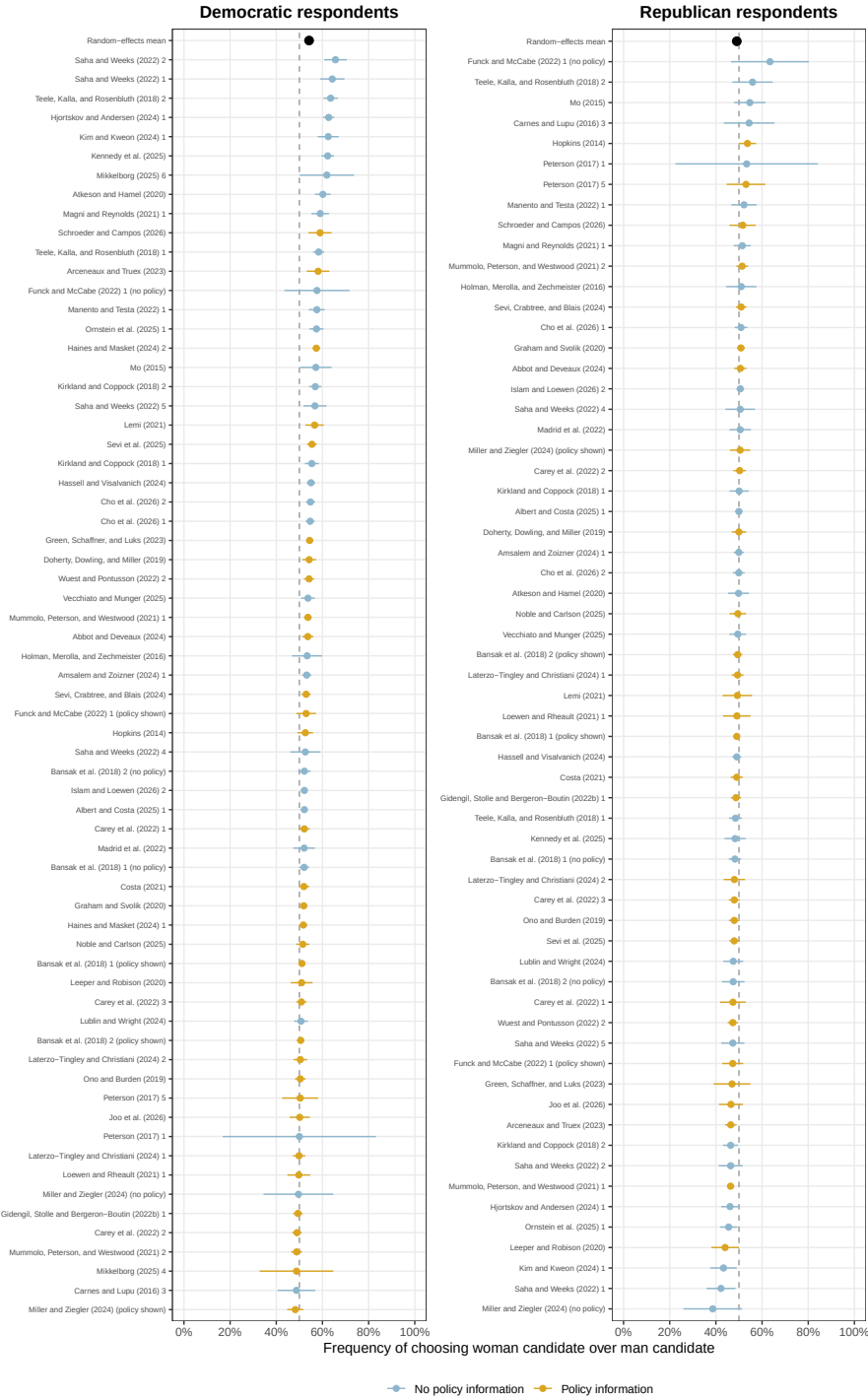


Figure O.24: Women versus Men win rates (US studies, by respondent party)
Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.



Figure O.25: Majority versus Minority win rates (international studies, part 1 of 2)
Note: Each row is one study estimate with its 95% confidence interval, for the first alphabetical half of the studies, grouped into US and non-US panels. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5. The majority-race category is defined relative to the country of study.

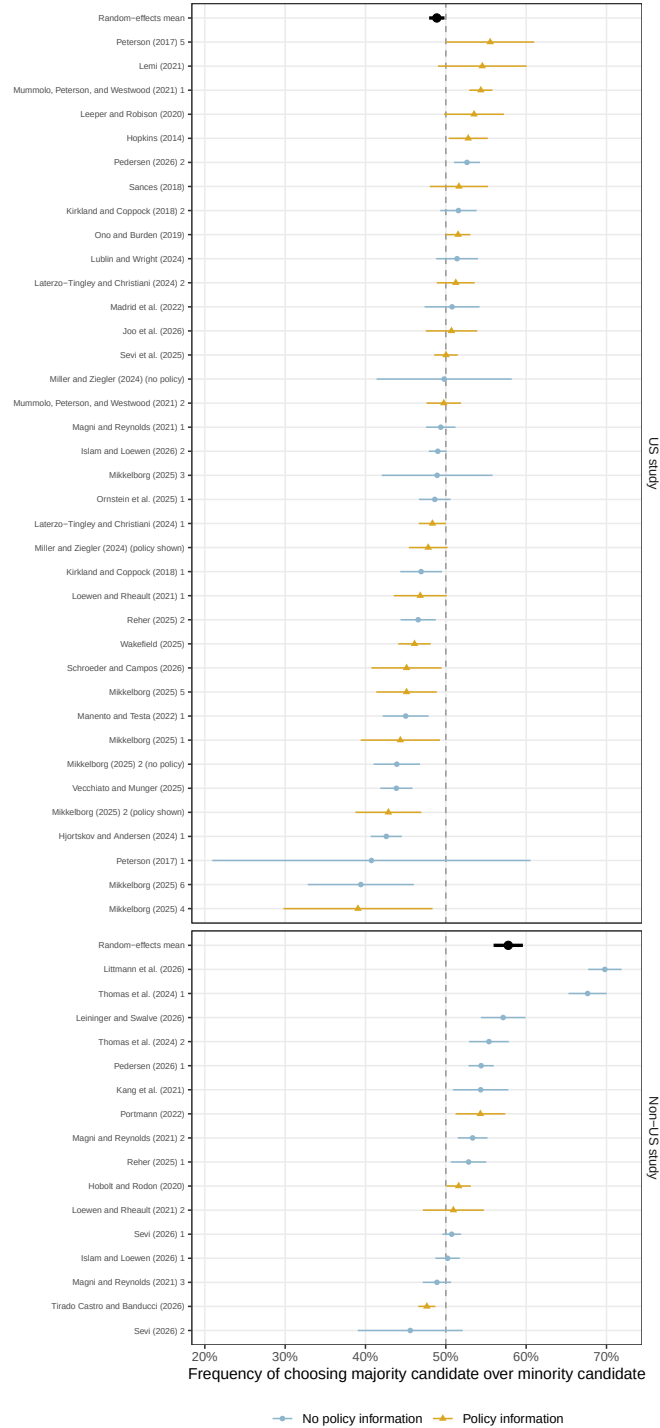


Figure O.26: Majority versus Minority win rates (international studies, part 2 of 2)
Note: Each row is one study estimate with its 95% confidence interval, for the second alphabetical half of the studies, grouped into US and non-US panels. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5. The majority-race category is defined relative to the country of study.

US race heterogeneity

The next figure isolates the white-versus-Black matchup in US studies and splits estimates by respondent race rather than respondent party. It anchors the main-text claim that pooled US race averages can mask offsetting subgroup tendencies.

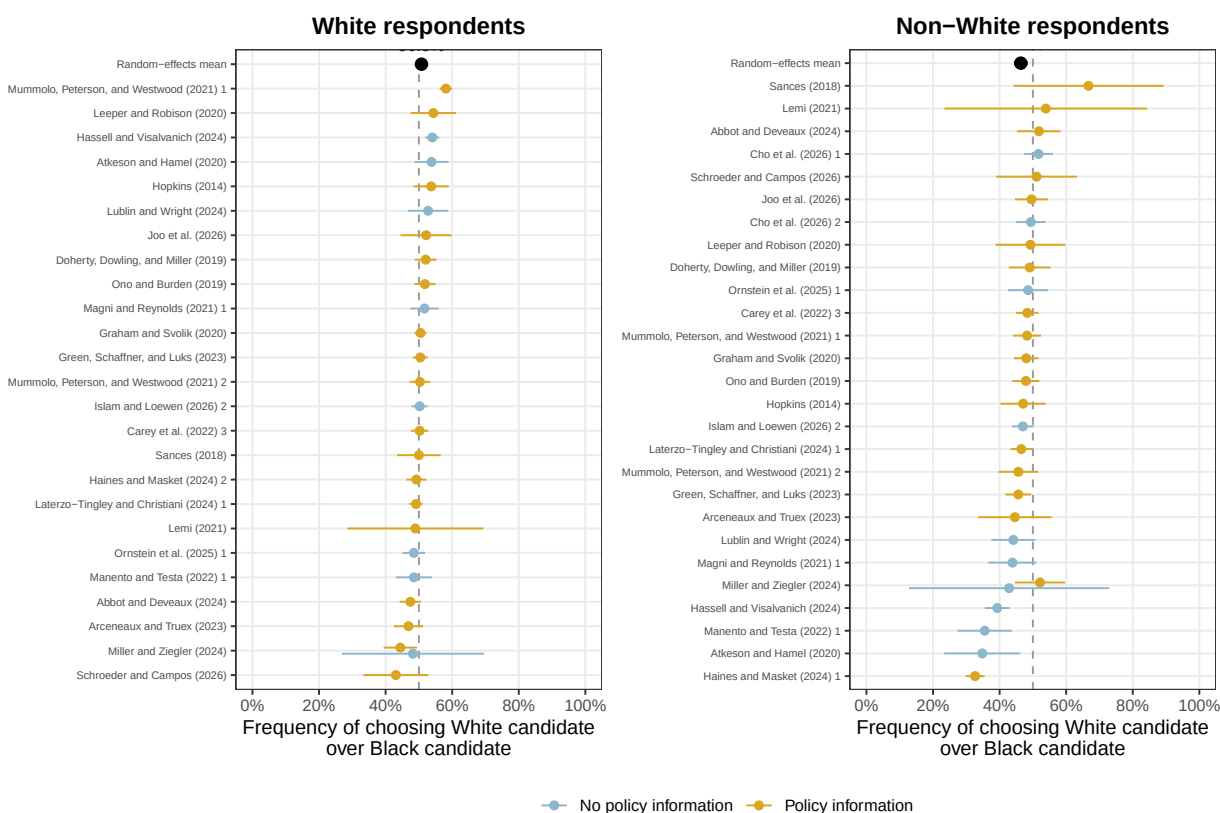


Figure O.27: White versus Black win rates (US studies, by respondent race)
Note: Each row is one study estimate, split into white and non-white respondent panels. The filled point and interval at top of each panel show $\hat{\mu}$ with its 95% confidence interval from the random-effects model. The dashed line marks 0.5. Gold points come from studies that include explicit policy information; blue points come from studies without. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

Table O.19 and Figure O.28 report the analysis closest to the shared-identification estimand of van Oosten, Mügge, and van der Pas (2024). Candidate A shares the respondent’s race or ethnicity; Candidate B does not. The first panel uses all eligible race-match tasks, while the second and third panels hold party-match status fixed among candidates who either both match or both do not match the respondent’s party.

Table O.19: Shared race/ethnicity candidate preference

Estimand	Respondents	k	μ [95% CI]	Premium over 50%
Shared race/ethnicity	All respondents	35	0.53 [0.52, 0.54]	3.0 pp
Shared race/ethnicity	White	34	0.51 [0.50, 0.52]	0.9 pp
Shared race/ethnicity	Black	33	0.59 [0.57, 0.61]	9.3 pp
Shared race/ethnicity	Hispanic	28	0.55 [0.53, 0.57]	5.4 pp
Shared race/ethnicity	Asian	18	0.58 [0.54, 0.62]	8.1 pp
Shared race/ethnicity, party match held present	All respondents	17	0.55 [0.51, 0.58]	4.7 pp
Shared race/ethnicity, party match held present	White	15	0.51 [0.49, 0.54]	1.5 pp
Shared race/ethnicity, party match held present	Black	14	0.64 [0.58, 0.69]	13.6 pp
Shared race/ethnicity, party match held present	Hispanic	13	0.53 [0.50, 0.57]	3.3 pp
Shared race/ethnicity, party match held present	Asian	8	0.63 [0.58, 0.69]	13.4 pp
Shared race/ethnicity, party match held absent	All respondents	21	0.55 [0.53, 0.58]	5.5 pp
Shared race/ethnicity, party match held absent	White	21	0.52 [0.50, 0.54]	2.4 pp
Shared race/ethnicity, party match held absent	Black	19	0.61 [0.57, 0.65]	10.9 pp
Shared race/ethnicity, party match held absent	Hispanic	17	0.58 [0.55, 0.62]	8.3 pp
Shared race/ethnicity, party match held absent	Asian	12	0.62 [0.58, 0.66]	11.8 pp

Rows report meta-analytic AFCPs for choosing the candidate who shares the respondent’s race/ethnicity over a candidate who does not. The first panel uses all eligible race-match tasks. The second and third panels hold party-match status constant when both candidates either do or do not match the respondent’s party. The analysis covers United States studies only, and among those is limited to studies with respondent and candidate race/ethnicity coded on compatible categories. Estimates are weighted to 2022 American Community Survey benchmarks.

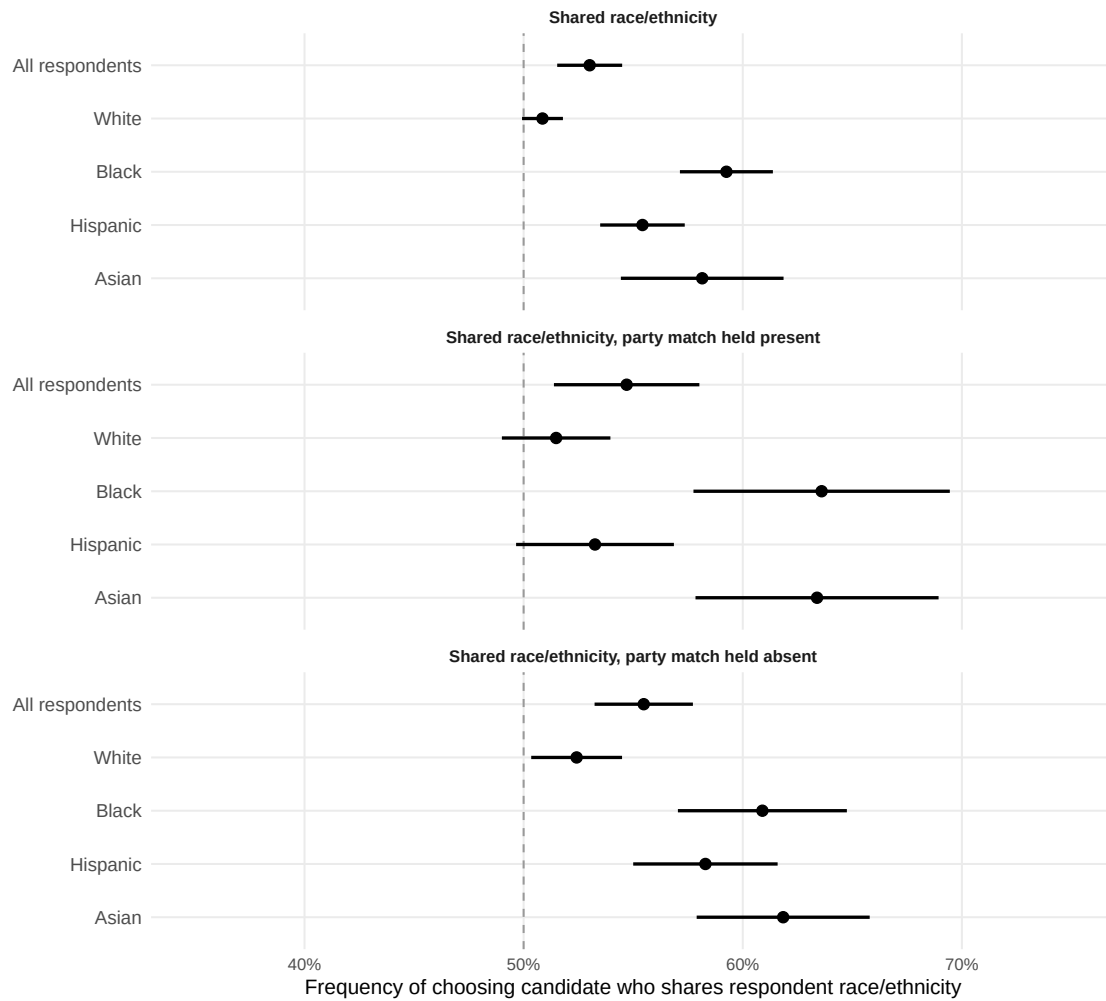


Figure O.28: Shared race/ethnicity candidate preference

Note: Estimates are restricted to US studies. Points are meta-analytic AFCP estimates; horizontal intervals are 95% confidence intervals. The dashed line marks 0.5. The analysis is restricted to studies in which respondent and candidate race or ethnicity can be coded on compatible categories.

The pooled shared-race estimate masks substantial variation across respondent groups. Figure O.29 disaggregates the shared-race preference by respondent race, estimating for each study the AFCEP of choosing a co-racial over a non-co-racial candidate among respondents of that same race. The co-racial premium is strong and consistent for Black respondents (+9.3 percentage points, 95% CI: 7.1–11.4, $k = 33$), moderate for Hispanic respondents (+5.4, 95% CI: 3.5–7.4, $k = 28$), positive but less precisely estimated for Asian respondents (+8.1, 95% CI: 4.4–11.9, $k = 18$), and essentially absent for white respondents (+0.9, 95% CI: –0.1 to 1.8, $k = 34$).

These magnitudes line up closely with the shared-identification meta-analysis of van Oosten, Mücke, and van der Pas (2024), who pool 43 candidate experiments and report that racial or ethnic minority voters favor co-ethnic candidates by 7.9 percentage points over majority candidates. Our minority-respondent estimates bracket their pooled figure, and like them we find no sharp ordering among the minority groups: although the Black premium is larger in point estimate than the Hispanic and Asian premiums, the confidence intervals overlap, echoing their result that the Both-Black, Both-Latinx, and Both-Asian summaries do not differ significantly. Our larger and more recent corpus, restricted to binary forced-choice tasks, reproduces their substantive conclusion that shared racial identification is a real and sizable driver of candidate choice for minority voters while being negligible for white voters.

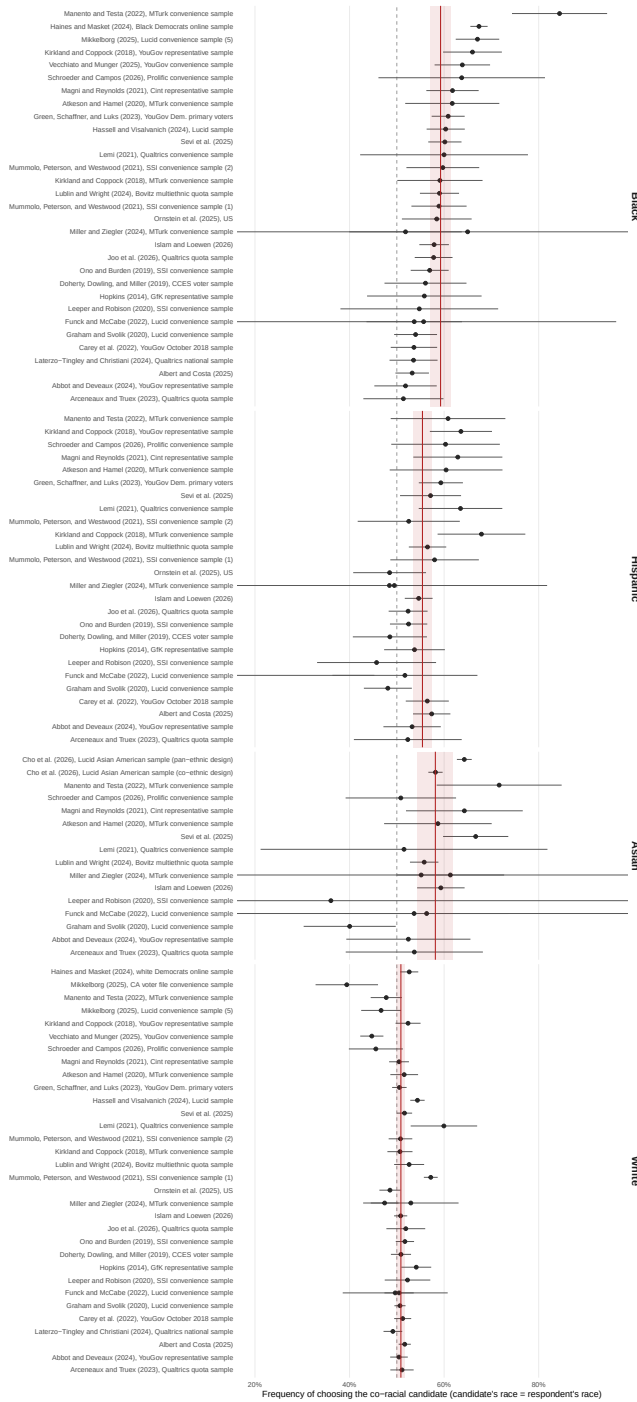


Figure O.29: Shared racial identification by study and respondent race
Note: Each point is a study-level AFCP of choosing the candidate whose race matches the respondent's, among respondents of that race, against a candidate of a different race. Intervals are 95% confidence intervals; the dashed line marks 0.5. The red line and shaded band give the random-effects meta-analytic mean and its 95% confidence interval within each respondent race. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

The shared-identification logic extends beyond race to every characteristic for which the corpus codes both a candidate attribute and the matching respondent attribute. The corresponding comparison across characteristics is reported in the main text’s “Shared Identification” section.

The following figure addresses the temporal-trend concern raised by Mikkelsen (2025). It restricts attention to white respondents in US white-versus-Black matchups and plots study-level estimates separately for Democrats and Republicans. In this broader corpus, support for the Black candidate among white Democrats rises by 0.8 percentage points per year, a slope distinguishable from zero ($p = 0.004$) that reproduces the trend Mikkelsen reports. Among white Republicans the slope is similar, 0.7 points per year ($p = 0.06$), and the two party slopes are not distinguishable from each other.

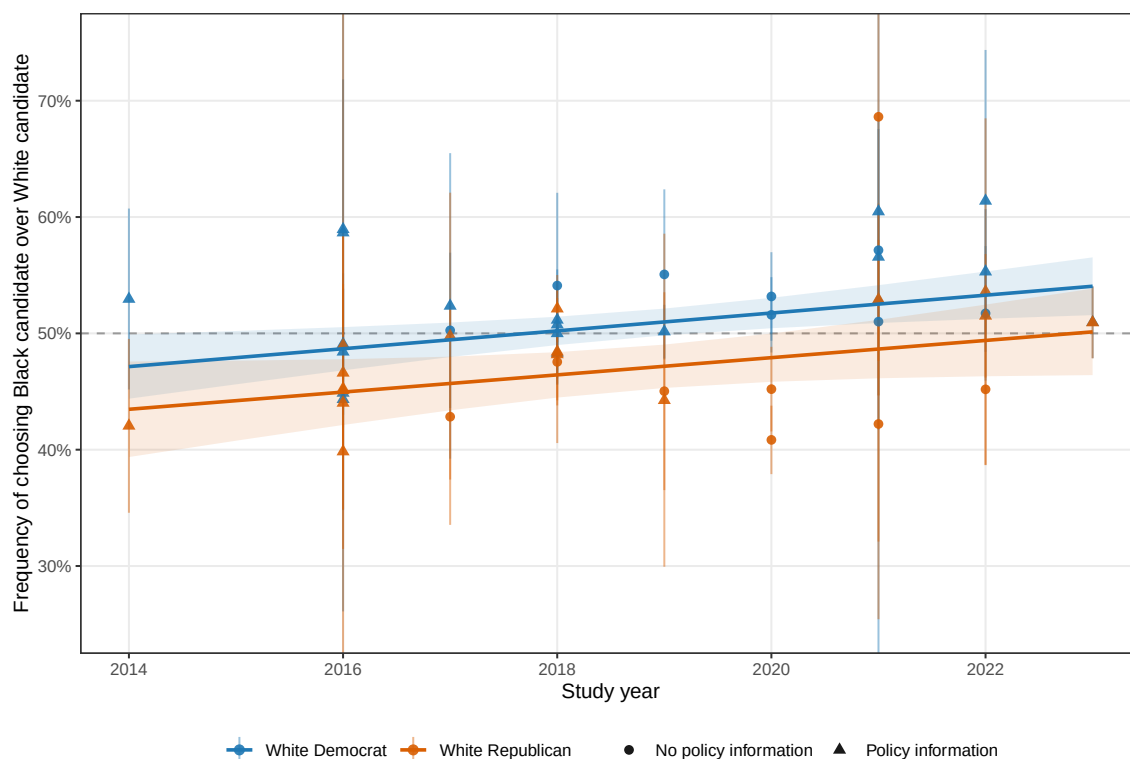


Figure O.30: Black versus white win rates over time among white respondents (US studies, by respondent party)

Note: Estimates are restricted to US studies. Each point is a study-level estimate of the frequency with which white respondents choose the Black candidate over the white candidate in direct white-versus-Black matchups. Vertical intervals are 95% confidence intervals. Solid lines and shaded bands are fitted values and 95% confidence intervals from separate random-effects meta-regressions by respondent party. Point shapes indicate whether the underlying study includes explicit policy information.

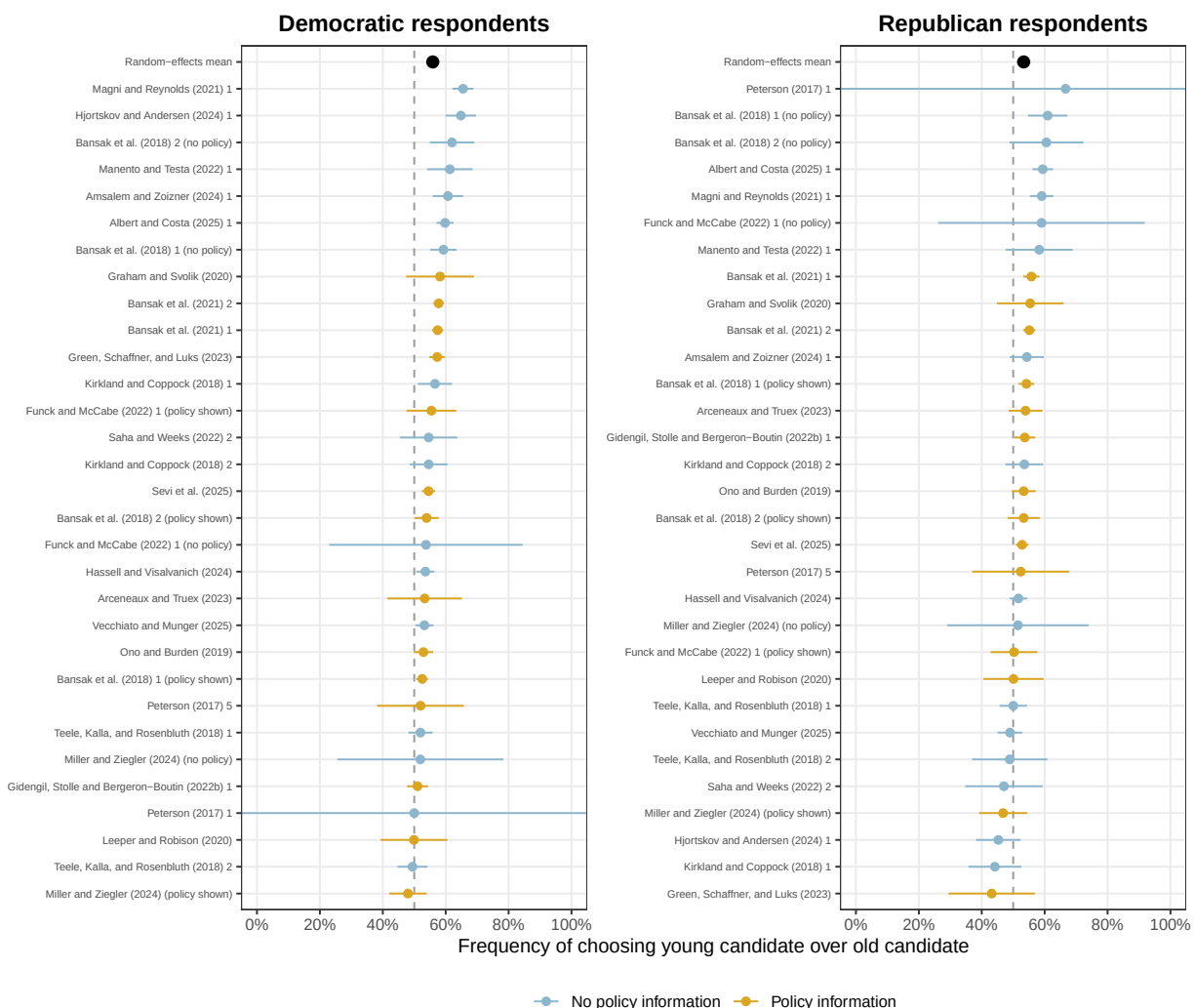


Figure O.31: Young versus Old win rates (US studies, by respondent party)
Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

Experience



Figure O.32: No versus Any Political Experience win rates (international studies, part 1 of 2)

Note: Each row is one study estimate with its 95% confidence interval, for the first alphabetical half of the studies, grouped into US and non-US panels. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5. An AFCP below 0.5 indicates that candidates without prior political experience are chosen less often than those with experience.

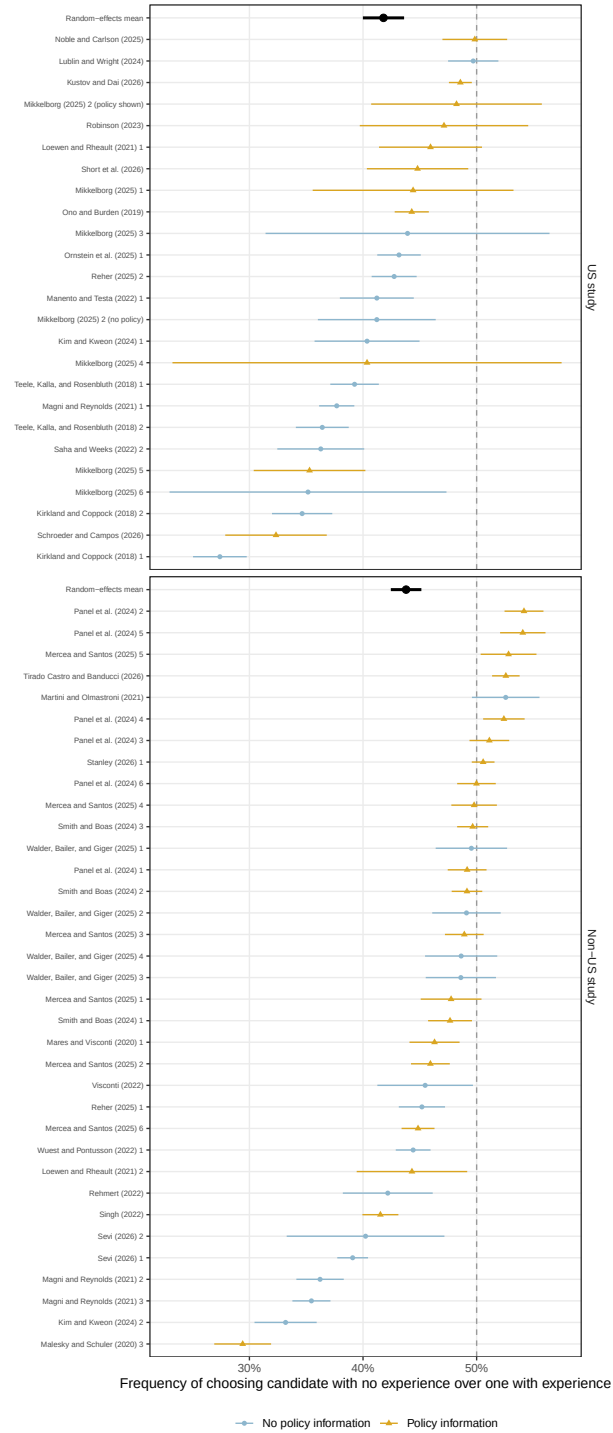


Figure O.33: No versus Any Political Experience win rates (international studies, part 2 of 2)

Note: Each row is one study estimate with its 95% confidence interval, for the second alphabetical half of the studies, grouped into US and non-US panels. Points are colored and shaped by whether the study shows policy information. The black point at the top of each panel is $\hat{\mu}$ with its 95% confidence interval from a random-effects model over *all* of that panel's studies, including the ones drawn in the companion figure. The dashed line marks 0.5. An AFCP below 0.5 indicates that candidates without prior political experience are chosen less often than those with experience.

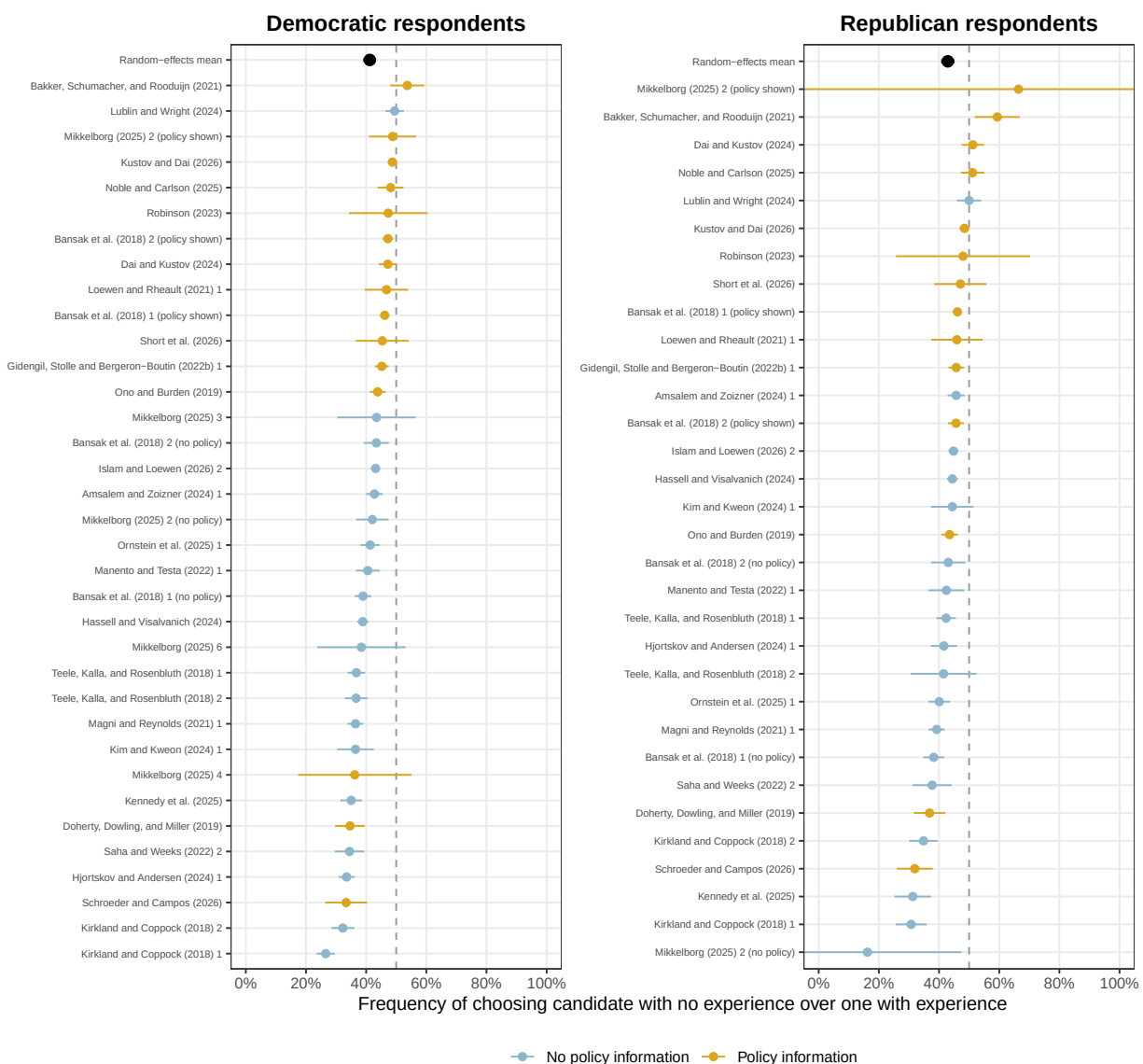


Figure O.34: No versus Any Political Experience win rates (US studies, by respondent party)
Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

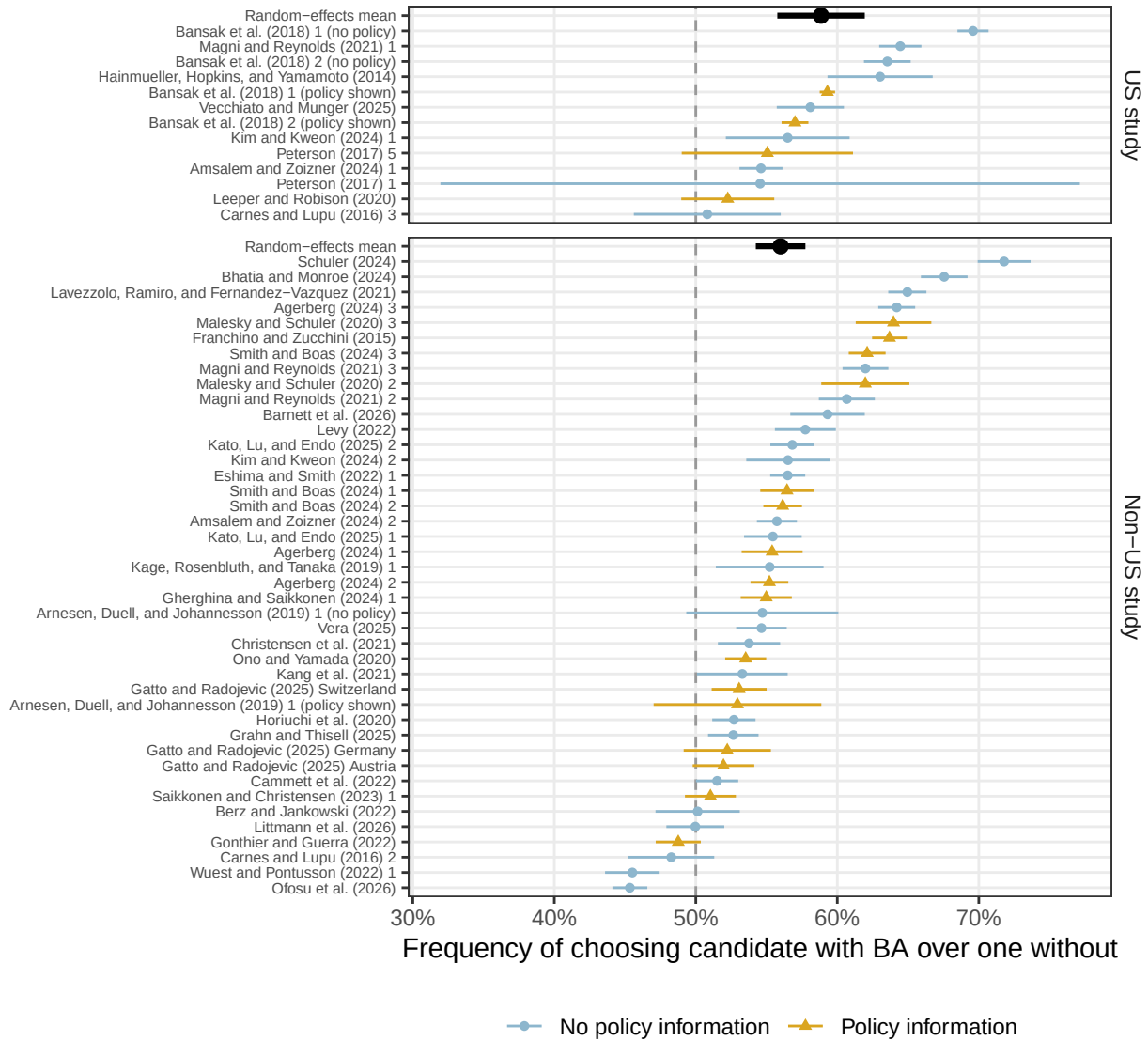


Figure O.35: College Degree versus None win rates (international studies)

Note: Estimates cover conjoints from all countries, not US studies only. Each row is one study estimate, in panels for US and for non-US studies. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. The dashed line marks 0.5. An AFCP above 0.5 indicates that candidates with a college degree are chosen more often than those without.

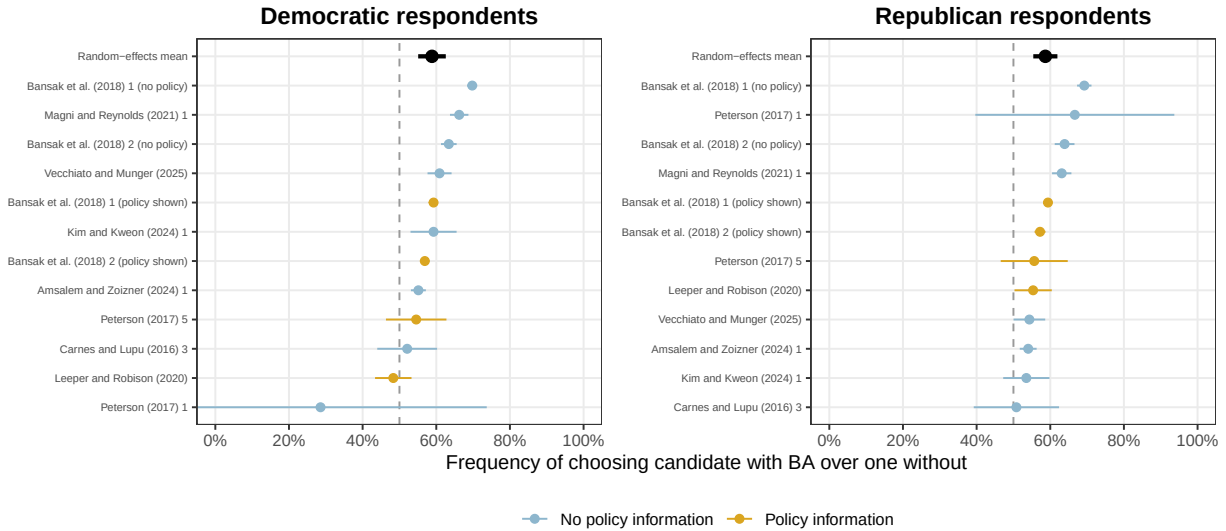


Figure O.36: College Degree versus None win rates (US studies, by respondent party)
Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

Politics

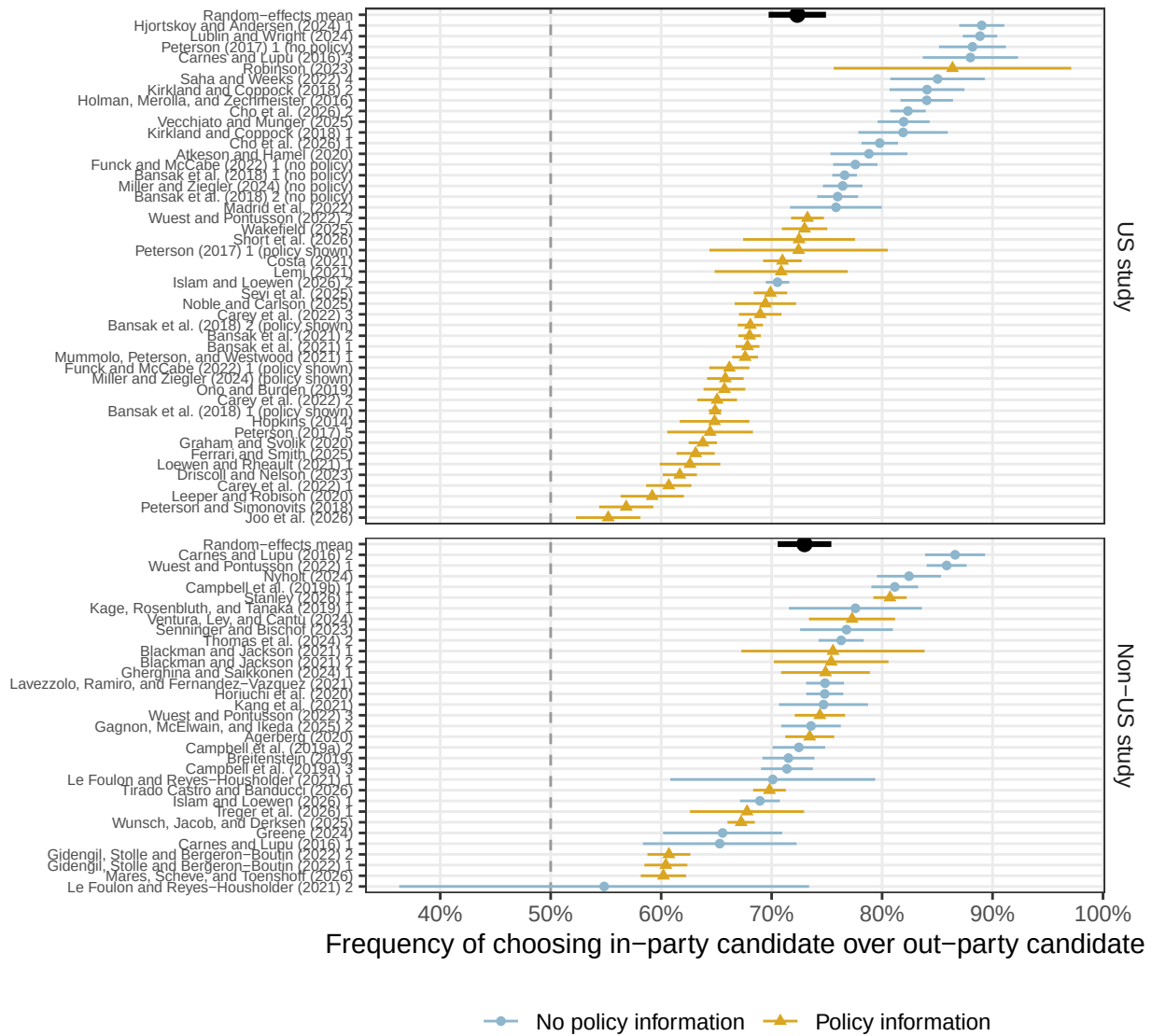


Figure O.37: In-party versus Out-party win rates (international studies)

Note: Estimates cover conjoints from all countries, not US studies only. Each row is one study estimate, in panels for US and for non-US studies. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. The dashed line marks 0.5. The AFCP is the probability of choosing the in-partisan candidate over the out-partisan candidate.

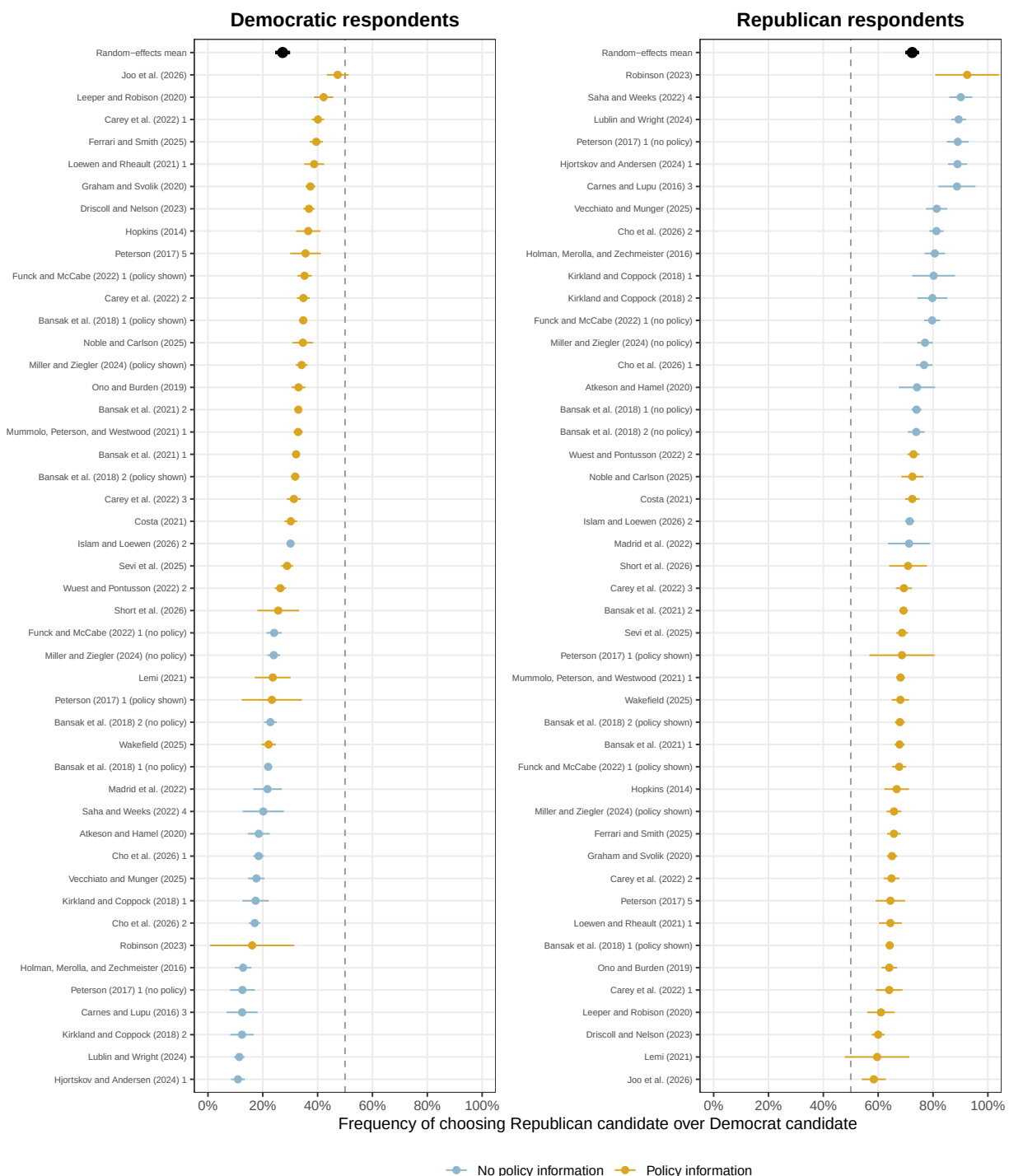


Figure O.38: Republican versus Democrat win rates (US studies, by respondent party)
Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

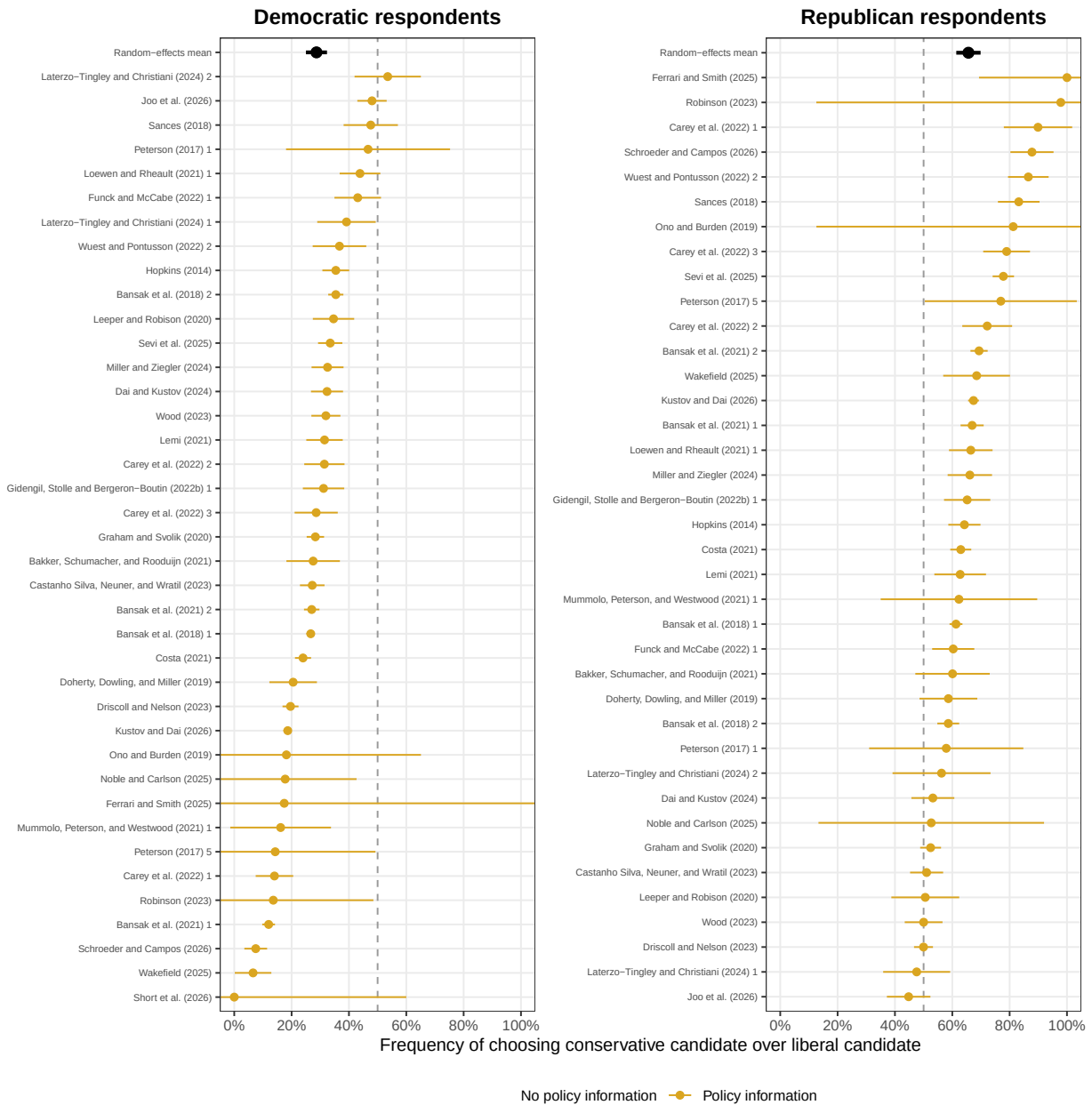


Figure O.39: Conservative versus Liberal policy positions win rates (US studies, by respondent party)

Note: Each row is one study estimate, in side-by-side panels for Democratic and for Republican respondents. The black point and interval at the top of each panel show $\hat{\mu}$ with its 95% confidence interval from a random-effects model over that panel's study estimates. Gold points come from studies that show policy information; blue points come from studies without. The dashed line marks 0.5. The AFCP below 0.5 for Democrats and above 0.5 for Republicans reflects partisan divergence in policy preferences. US respondents are weighted to 2022 ACS benchmarks where the replication data permit and are otherwise unweighted.

P Covariate-adjusted AFCPs

The AFCP for levels A and B of an attribute is estimated on the tasks that pit an A profile against a B profile. Random assignment ensures that, in expectation, the A and B profiles in those tasks are alike on every other attribute. In any one study they will differ somewhat by chance, and the raw AFCP then reflects both the A-versus-B contrast and whatever the realized imbalance happens to favor. Because the other attributes are themselves randomized, adjusting for them is justified by the design, not by a modeling assumption about the outcome. The adjustment removes the chance component.

For a matchup between levels A and B, let $Y_i = 1$ if the A profile in task i was chosen. For every other attribute in the study, we code indicators for each level except a reference level and take the difference between the two profiles, $D_i = X_{A,i} - X_{B,i}$. We then estimate

$$Y_i = \alpha + D_i' \beta + \varepsilon_i$$

by least squares with the same survey weights and the same respondent-clustered (CR2) standard errors as the raw estimates. The adjusted AFCP is $\hat{\alpha}$.

The centering of D_i determines what α means. If D_i were centered at its sample mean, $\hat{\alpha}$ would equal the sample mean of Y_i by the algebra of least squares, so the “adjusted” estimate would be the raw AFCP exactly. We instead center D_i at its expectation under the design. That expectation is zero whenever an attribute is assigned independently of the focal one, because the A and B profiles are then exchangeable with respect to it. Leaving D_i uncentered therefore centers it at the right place. In other words, α is the predicted probability that the A profile wins a task in which the two profiles differ on nothing else, which under additivity in β is the AFCP with the realized imbalance removed.

Three kinds of attributes are kept out of the adjustment. First, an attribute tied to the focal attribute by the design has a nonzero expected difference, and centering it at zero would extrapolate to profiles the design never shows. We detect such ties in two ways. An attribute is dropped when some combination of its levels with the focal levels never appears among the profiles even though independent assignment would produce ten or more, or when a chi-square test rejects independence between the attribute and the focal attribute at the 0.01 level. The first rule catches the raw column from which a standardized attribute was coded and the issue positions from which an overall policy profile is built. The second catches restrictions that our coarser categories hide: Pedersen (2026) gives 25-to-34-year-old candidates one to three years of experience, but our youngest age category also holds

40-to-44-year-olds with any amount of experience, so no combination is empty even though the young candidates in a young-versus-middle-aged task have systematically less experience than their opponents. Dropping an attribute that is in fact independent costs a little precision, while keeping a linked one moves the estimate, so the threshold is a liberal one. For the in-party matchup we also drop the party labels, since in-party status is built from them. Second, we require at least 20 tasks per adjustment term and at least five distinct configurations of D_i per estimated parameter, shedding the attributes with the most levels first. The second requirement binds in fixed or fractional designs, where a matchup can realize only a handful of profile pairings and $\hat{\alpha}$ would be an extrapolation to a pairing no respondent saw. These requirements leave 183 of the 1,815 estimates unadjusted. Because each column of D_i takes only the values -1 , 0 , and 1 , k terms can realize at most 3^k configurations, so no adjustment with fewer than three terms can meet the configuration requirement, and three terms need 60 tasks. Fifty-two of the 183 estimates rest on fewer than 60 tasks; the other 131 have the tasks but too few distinct configurations. Third, one study records a respondent’s rating of each profile alongside the attributes. Because that rating is measured after treatment, we exclude it.

Figure P.40 plots the adjusted estimates against the raw ones. The two agree closely (correlation 0.998). The median absolute change is 0.18 percentage points and the 90th percentile is 0.85; 93% of estimates move by less than one point, and 93% by less than half of their raw standard error. Forty-six estimates (2.5%) cross 0.5. All were within 2.4 points of 0.5 to begin with, and none was distinguishable from it. The largest change, 8.1 points, belongs to the white-versus-Black comparison in Lemi (2021), estimated on 172 tasks with a raw standard error of 7.4 points; the other labeled points are likewise small-sample estimates. The adjustment buys little precision (the median ratio of adjusted to raw standard errors is 0.98). In short, chance imbalance in the other attributes does not drive any of the estimates we report, and the main text uses the raw AFCPs.

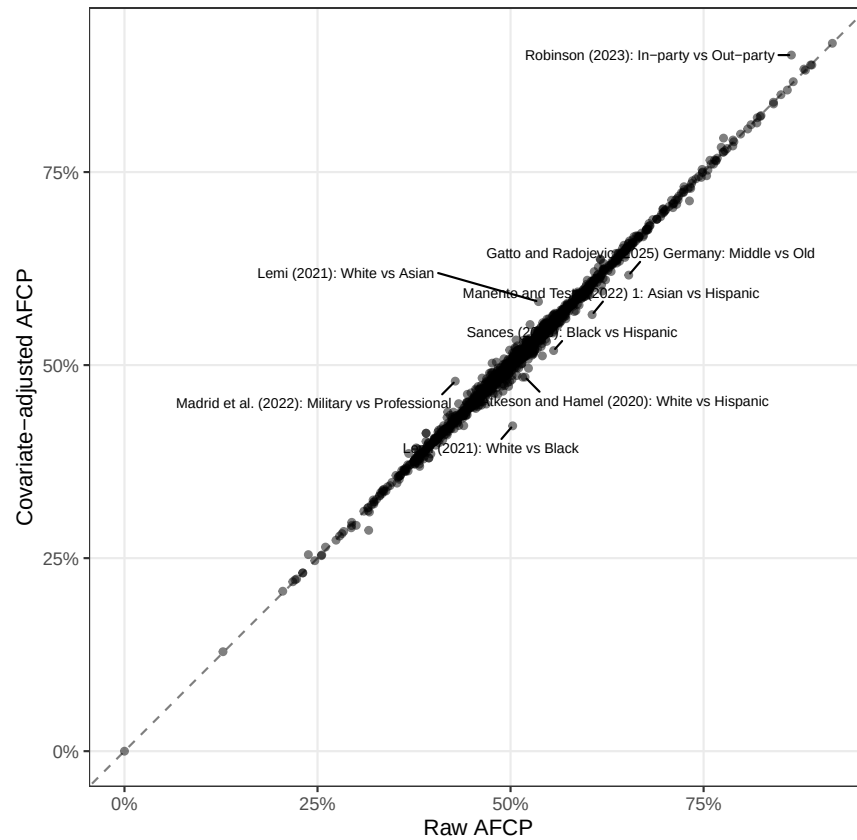


Figure P.40: Raw and covariate-adjusted AFCPs

Note: Each point is one study-level AFCP (one per policy condition where a study randomizes whether policy is shown), 1,630 in all. The adjusted estimate is the intercept from a regression of the A-profile win indicator on the A-minus-B differences in the study's other randomized attributes, uncentered; see the text for the attributes excluded and the support requirements. The dashed line is the 45-degree line. The eight largest changes are labeled.

Q Estimates underlying the main-text figures

Every estimate reported in this paper comes from a randomized comparison within a study, pooled across studies by random-effects meta-analysis. No estimate in the main text or in these tables adjusts for covariates, so there are no control coefficients to report; the covariate-adjusted comparison in Appendix P reports only the adjusted intercepts. The tables in this section give the complete numerical results behind each main-text figure that reports estimates: the pooled Average Feature Choice Probability $\hat{\mu}$, its standard error and 95% confidence interval, the estimated between-study standard deviation $\hat{\tau}$, the p -value of the heterogeneity test, the number of study-level estimates pooled, and the sample sizes in respondents and choice tasks. The dependent variable throughout is `cand.selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task.

One estimation detail affects a small number of study-level estimates. A comparison that is available to only a handful of respondents within a study can produce a forced-choice probability of exactly 0 or exactly 1, and the usual sampling variance of such an estimate is zero. An estimate with no sampling variance cannot enter an inverse-variance weighted meta-analysis at all: dropping it removes the study from the pooled k while its point remains on the plot, and retaining it gives one study unbounded weight. We take the standard error instead from the Agresti–Coull adjusted proportion (Agresti and Coull, 1998; Gelman and Hill, 2007), which adds one pseudo-success and one pseudo-failure before computing it: with x choices of candidate A among n respondents, $\tilde{p} = (x + 1)/(n + 2)$ and the standard error is $\sqrt{\tilde{p}(1 - \tilde{p})/n}$. The point estimate is left at its observed value, so every point we plot is what the study found; only the confidence interval and the meta-analytic weight come from the adjustment, and the interval a figure draws is therefore the same one the pooling uses. We take n to be the number of respondents rather than the number of tasks, because our standard errors are clustered by respondent and a task count would credit the comparison with more information than the design supplies. This adjustment applies to 59 of the 6,828 study-level estimates computed anywhere in the paper (0.9%; the count spans the full-sample, respondent-subgroup, affinity, party–policy, and shared-identification analyses), spread across 22 studies and each resting on between 2 and 12 respondents. Because the adjusted standard errors are large, these estimates receive little weight wherever they enter; the point of the adjustment is that they enter at all, and that a reader counting points on a figure arrives at the k printed beside it.

Figures 1, 2, and 3 of the main text are descriptive rather than inferential: they display which attributes each study randomized and how many tasks each presented to a

respondent. The study-by-study task counts behind them are in the study manifest, Table C.2; that table carries no attribute count, and the attribute counts in Figure 1 come from `meta_data/conjoint_studies.xlsx` in the replication archive.

Each table below reports the pooled estimates the corresponding figure plots. The study-level estimates pooled into them are reported in full in a separate estimate-level supplement, which typesets the 6,220 study-level estimates behind the main-text figures, from 239 studies, as Tables S1 through S7 (an estimate that a figure plots twice appears twice), one table per main-text figure, one row per study-level estimate, grouped by the comparison the figure draws. Each table below names its counterpart there. The same content is also deposited in machine-readable form in the replication archive as `output/tables/supplement_estimate_level.csv`, with a README documenting the columns.

Table Q.20: Estimates underlying Figure 4 (international matchups)

Matchup	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	Studies	N resp.	N tasks
Women vs Men	No policy information	0.536	0.003	[0.530, 0.542]	0.029	0.000	119	119	119	142,710	364,158
Women vs Men	Policy included	0.523	0.005	[0.514, 0.533]	0.046	0.000	99	99	99	172,724	480,132
Women vs Men	Overall	0.530	0.003	[0.525, 0.536]	0.039	0.000	218	212	212	315,434	844,290
Young vs Old	No policy information	0.588	0.007	[0.573, 0.602]	0.054	0.000	63	63	63	39,432	64,295
Young vs Old	Policy included	0.559	0.005	[0.550, 0.568]	0.028	0.000	51	51	51	62,191	142,747
Young vs Old	Overall	0.575	0.005	[0.566, 0.585]	0.048	0.000	114	109	109	101,623	207,042
Young vs Middle	No policy information	0.508	0.004	[0.501, 0.515]	0.029	0.000	80	80	80	74,735	139,567
Young vs Middle	Policy included	0.515	0.004	[0.508, 0.523]	0.024	0.000	53	53	53	93,036	215,946
Young vs Middle	Overall	0.511	0.003	[0.506, 0.516]	0.027	0.000	133	126	126	167,771	355,513
Middle vs Old	No policy information	0.597	0.007	[0.584, 0.610]	0.046	0.000	61	61	61	45,135	80,840
Middle vs Old	Policy included	0.549	0.006	[0.536, 0.561]	0.032	0.000	34	34	34	42,572	101,303
Middle vs Old	Overall	0.580	0.005	[0.570, 0.591]	0.049	0.000	95	90	90	87,707	182,143
Majority vs Minority	No policy information	0.542	0.010	[0.523, 0.560]	0.075	0.000	64	64	64	67,267	154,814
Majority vs Minority	Policy included	0.502	0.005	[0.492, 0.513]	0.033	0.000	42	42	42	49,869	148,465
Majority vs Minority	Overall	0.526	0.006	[0.513, 0.538]	0.065	0.000	106	100	100	117,136	303,279
Christian vs None	No policy information	0.487	0.015	[0.458, 0.516]	0.044	0.000	12	12	12	6,271	15,457
Christian vs None	Policy included	0.492	0.009	[0.475, 0.509]	0.021	0.001	10	10	10	7,776	37,189
Christian vs None	Overall	0.488	0.009	[0.471, 0.505]	0.034	0.000	22	16	16	14,047	52,646
Christian vs Muslim	No policy information	0.616	0.015	[0.586, 0.645]	0.031	0.000	6	6	6	4,315	6,730
Christian vs Muslim	Policy included	0.573	0.014	[0.546, 0.600]	0.000	0.574	3	3	3	1,242	1,966
Christian vs Muslim	Overall	0.606	0.013	[0.581, 0.631]	0.031	0.000	9	7	7	5,557	8,696
Muslim vs None	No policy information	0.368	0.015	[0.339, 0.397]	0.029	0.000	6	6	6	4,819	7,822

Table Q.20: Estimates underlying Figure 4 (international matchups) (*continued*)

Matchup	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	Studies	N resp.	N tasks
Muslim vs None	Policy included	0.378	0.013	[0.353, 0.403]	0.001	0.319		3	3	1,064	1,740
Muslim vs None	Overall	0.370	0.012	[0.347, 0.393]	0.026	0.000		9	7	5,883	9,562
College Degree vs None	No policy information	0.570	0.011	[0.548, 0.592]	0.064	0.000		35	35	48,008	136,285
College Degree vs None	Policy included	0.559	0.010	[0.540, 0.579]	0.042	0.000		20	20	34,736	153,047
College Degree vs None	Overall	0.566	0.008	[0.551, 0.582]	0.057	0.000		55	52	82,744	289,332
No vs Any Political Experience	No policy information	0.407	0.006	[0.395, 0.420]	0.051	0.000		72	72	75,376	165,733
No vs Any Political Experience	Policy included	0.468	0.008	[0.453, 0.483]	0.050	0.000		48	48	66,992	188,394
No vs Any Political Experience	Overall	0.432	0.006	[0.421, 0.443]	0.059	0.000		120	116	142,368	354,127
Business/Managerial vs Professional	No policy information	0.488	0.005	[0.479, 0.497]	0.028	0.000		62	62	35,071	54,020
Business/Managerial vs Professional	Policy included	0.478	0.006	[0.466, 0.489]	0.034	0.000		48	48	34,391	51,355
Business/Managerial vs Professional	Overall	0.484	0.004	[0.476, 0.491]	0.031	0.000		110	103	69,462	105,375
Educator vs Professional	No policy information	0.492	0.007	[0.479, 0.505]	0.039	0.000		53	53	26,703	36,595
Educator vs Professional	Policy included	0.505	0.005	[0.494, 0.516]	0.012	0.129		22	22	13,738	17,628
Educator vs Professional	Overall	0.496	0.005	[0.486, 0.505]	0.033	0.000		75	72	40,441	54,223
Farmer vs Professional	No policy information	0.509	0.010	[0.489, 0.529]	0.034	0.000		21	21	8,518	12,298
Farmer vs Professional	Policy included	NA	NA	[NA, NA]	NA	NA	NA		14	8,078	15,930
Farmer vs Professional	Overall	0.508	0.006	[0.497, 0.520]	0.022	0.000		35	30	16,596	28,228
Military vs Professional	No policy information	0.495	0.024	[0.448, 0.542]	0.042	0.005		6	6	2,765	3,703
Military vs Professional	Policy included	0.487	0.005	[0.477, 0.497]	0.002	0.386		8	8	7,396	10,547
Military vs Professional	Overall	0.487	0.007	[0.473, 0.502]	0.018	0.019		14	12	10,161	14,250
Working Class vs Professional	No policy information	0.476	0.014	[0.448, 0.504]	0.057	0.000		19	19	14,408	22,698
Working Class vs Professional	Policy included	0.502	0.006	[0.491, 0.513]	0.016	0.000		16	16	17,911	31,665
Working Class vs Professional	Overall	0.489	0.009	[0.472, 0.506]	0.045	0.000		35	31	32,319	54,363
In-party vs Out-party	No policy information	0.781	0.011	[0.760, 0.802]	0.063	0.000		37	37	32,502	112,373
In-party vs Out-party	Policy included	0.677	0.009	[0.659, 0.695]	0.058	0.000		42	42	47,640	266,769

In-party vs Out-party	Overall	0.725	0.009	[0.707, 0.744]	0.081	0.000	79	74	80,142	379,142
-----------------------	---------	-------	-------	----------------	-------	-------	----	----	--------	---------

Note: Estimates cover conjoints from all countries, not US studies only. Full numerical results for every point plotted in Figure 4 of the main text. The dependent variable throughout is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task; $\hat{\mu}$ is the meta-analytic average forced-choice probability (AFCP) that the first-named profile is chosen, so 0.500 is an even split. Estimates come from randomized comparisons within study and are pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and Q p the p -value of the heterogeneity test. k counts the study-level estimates pooled and N the respondents and choice tasks behind them; a study that randomizes whether policy information is shown contributes one estimate per condition, so k exceeds the study count reported alongside it. Study estimates with an observed sampling variance of zero take their standard error from the Agresti–Coull adjustment described in the text of this section and are pooled, and counted in k and N , like any other. Standard errors within study are clustered by respondent. Every study-level estimate pooled into a row here appears, one row per estimate, in the estimate-level supplement, Table S1, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

Table Q.21: Estimates underlying Figure 5 (US matchups, by respondent partisanship)

Matchup	Respondents	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	Studies	N resp.	N tasks
Women vs Men	Republicans	0.491	0.002	[0.486, 0.495]	0.010	0.007	62	58	25,555	90,515	
Women vs Men	Democrats	0.542	0.005	[0.532, 0.552]	0.037	0.000	66	62	39,106	147,146	
Young vs Old	Republicans	0.533	0.007	[0.519, 0.548]	0.027	0.000	31	27	9,755	22,565	
Young vs Old	Democrats	0.559	0.009	[0.542, 0.576]	0.038	0.000	31	27	15,009	33,982	
Young vs Middle	Republicans	0.495	0.004	[0.488, 0.502]	0.009	0.066	35	30	13,257	35,524	
Young vs Middle	Democrats	0.500	0.006	[0.489, 0.512]	0.027	0.000	38	33	20,987	56,172	
Middle vs Old	Republicans	0.558	0.005	[0.548, 0.569]	0.016	0.001	29	25	10,843	28,365	
Middle vs Old	Democrats	0.562	0.007	[0.549, 0.574]	0.027	0.000	29	25	16,864	46,726	
White vs Black	Republicans	0.522	0.006	[0.510, 0.533]	0.027	0.000	50	45	13,005	25,049	
White vs Black	Democrats	0.466	0.007	[0.453, 0.479]	0.041	0.000	55	50	23,573	43,727	
White vs Hispanic	Republicans	0.530	0.006	[0.518, 0.542]	0.026	0.000	44	40	11,811	21,989	
White vs Hispanic	Democrats	0.487	0.005	[0.478, 0.497]	0.021	0.000	46	42	18,727	35,200	
White vs Asian	Republicans	0.515	0.010	[0.496, 0.534]	0.040	0.000	32	28	7,341	14,256	
White vs Asian	Democrats	0.482	0.010	[0.463, 0.502]	0.047	0.000	34	30	11,584	23,919	
Asian vs Hispanic	Republicans	0.541	0.011	[0.520, 0.562]	0.043	0.000	30	27	6,194	10,978	
Asian vs Hispanic	Democrats	0.508	0.010	[0.489, 0.527]	0.043	0.000	33	29	10,062	18,759	
Black vs Hispanic	Republicans	0.501	0.005	[0.491, 0.511]	0.008	0.380	42	38	8,489	14,815	
Black vs Hispanic	Democrats	0.520	0.007	[0.506, 0.535]	0.034	0.000	44	40	13,798	24,104	
Christian vs None	Republicans	0.535	0.008	[0.519, 0.551]	0.016	0.060	13	9	2,576	14,240	
Christian vs None	Democrats	0.484	0.007	[0.471, 0.497]	0.013	0.135	13	9	4,075	26,307	
Christian vs Muslim	Republicans	0.744	0.020	[0.705, 0.783]	0.001	0.269	3	3	504	742	
Christian vs Muslim	Democrats	0.572	0.025	[0.524, 0.621]	0.027	0.218	3	3	674	1,068	
Muslim vs None	Republicans	0.338	0.024	[0.291, 0.385]	0.000	0.745	3	3	372	463	
Muslim vs None	Democrats	0.390	0.065	[0.264, 0.517]	0.097	0.016	3	3	459	579	
College Degree vs None	Republicans	0.587	0.017	[0.554, 0.619]	0.051	0.000	12	10	4,704	52,371	
College Degree vs None	Democrats	0.588	0.019	[0.551, 0.626]	0.060	0.000	12	10	7,084	90,713	
No vs Any Political Experience	Republicans	0.429	0.011	[0.407, 0.451]	0.056	0.000	32	29	11,764	40,434	
No vs Any Political Experience	Democrats	0.412	0.011	[0.391, 0.433]	0.058	0.000	35	32	18,530	63,778	
Business/Managerial vs Professional	Republicans	0.531	0.009	[0.513, 0.549]	0.034	0.000	31	27	6,687	11,684	
Business/Managerial vs Professional	Democrats	0.491	0.007	[0.477, 0.506]	0.027	0.000	34	29	9,609	18,131	
Educator vs Professional	Republicans	0.510	0.014	[0.482, 0.537]	0.048	0.004	27	24	4,343	5,651	
Educator vs Professional	Democrats	0.518	0.011	[0.496, 0.539]	0.039	0.000	29	26	6,474	8,380	
Farmer vs Professional	Republicans	0.507	0.010	[0.488, 0.527]	0.020	0.326	18	14	3,537	7,085	
Farmer vs Professional	Democrats	0.498	0.005	[0.488, 0.509]	0.004	0.046	18	14	4,663	10,970	
Military vs Professional	Republicans	0.529	0.027	[0.476, 0.583]	0.048	0.058	8	6	946	1,149	
Military vs Professional	Democrats	0.492	0.015	[0.463, 0.521]	0.001	0.707	8	6	1,636	2,028	
Working Class vs Professional	Republicans	0.491	0.019	[0.455, 0.527]	0.055	0.000	13	11	3,285	7,090	
Working Class vs Professional	Democrats	0.487	0.007	[0.473, 0.500]	0.011	0.297	13	11	4,785	11,463	
Conservative vs Liberal Issue Positions	Republicans	0.656	0.022	[0.614, 0.699]	0.121	0.000	38	38	9,854	30,499	
Conservative vs Liberal Issue Positions	Democrats	0.287	0.019	[0.250, 0.323]	0.105	0.000	39	39	14,161	46,435	

Table Q.21: Estimates underlying Figure 5 (US matchups, by respondent partisanship)
(*continued*)

Matchup	Respondents	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	Studies	N resp.	N tasks
Conservative vs Mixed Issue Positions	Republicans	0.599	0.015	[0.570, 0.628]	0.080	0.000	36	36	36	11,227	31,186
Conservative vs Mixed Issue Positions	Democrats	0.322	0.018	[0.286, 0.358]	0.105	0.000	36	36	36	15,522	46,875
Liberal vs Mixed Issue Positions	Republicans	0.359	0.018	[0.323, 0.394]	0.102	0.000	36	36	36	10,683	29,144
Liberal vs Mixed Issue Positions	Democrats	0.654	0.017	[0.621, 0.688]	0.096	0.000	36	36	36	15,487	46,499
Republican vs Democrat	Republicans	0.724	0.013	[0.699, 0.749]	0.085	0.000	46	41	41	21,136	121,057
Republican vs Democrat	Democrats	0.272	0.014	[0.245, 0.300]	0.093	0.000	46	41	41	30,409	190,822
Republican vs Independent	Republicans	0.597	0.023	[0.552, 0.642]	0.039	0.110	7	7	7	994	1,603
Republican vs Independent	Democrats	0.249	0.029	[0.193, 0.305]	0.070	0.000	7	7	7	1,939	3,265
Democrat vs Independent	Republicans	0.313	0.032	[0.250, 0.375]	0.072	0.000	7	7	7	974	1,622
Democrat vs Independent	Democrats	0.596	0.015	[0.567, 0.625]	0.021	0.229	7	7	7	1,832	3,073

Note: Full numerical results for every point plotted in Figure 5 of the main text. Rows are restricted to US studies. The dependent variable throughout is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task; $\hat{\mu}$ is the meta-analytic average forced-choice probability (AFCP) that the first-named profile is chosen, so 0.500 is an even split. Estimates are conditional on respondent party identification, computed within study over randomized comparisons and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and Q p the p -value of the heterogeneity test. k counts the study-level estimates pooled and N the choice tasks behind them; a study that randomizes whether policy information is shown contributes one estimate per condition, so k can exceed the study count reported alongside it. Study estimates with an observed sampling variance of zero take their standard error from the Agresti–Coull adjustment described in the text of this section and are pooled, and counted in k and N , like any other. Standard errors within study are clustered by respondent. Study-level estimates behind each row are in the estimate-level supplement, Table S2, and in machine-readable form at `output/tables/supplement_estimate_level.csv`, one row per estimate.

Table Q.22: Estimates underlying Figure 6 (shared identification)

Characteristic	Subsample	$\hat{\mu}$	SE	95% CI	Premium (pp)	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Age	No policy information	0.524	0.004	[0.517, 0.531]	+2.4	0.017	0.000		38	47,048	108,232
Age	Overall	0.521	0.002	[0.517, 0.525]	+2.1	0.015	0.000		89	118,240	336,832
Age	Policy included	0.519	0.002	[0.514, 0.524]	+1.9	0.013	0.000		51	71,192	228,600
Education	No policy information	0.497	0.010	[0.477, 0.517]	-0.3	0.051	0.000		26	37,446	108,992
Education	Overall	0.500	0.008	[0.485, 0.514]	-0.0	0.045	0.000		38	60,592	239,646
Education	Policy included	0.505	0.008	[0.489, 0.521]	+0.5	0.026	0.000		12	23,146	130,654
Gender	No policy information	0.518	0.002	[0.515, 0.522]	+1.8	0.013	0.000		109	132,545	332,754
Gender	Overall	0.515	0.001	[0.513, 0.517]	+1.5	0.011	0.000		196	273,683	709,120
Gender	Policy included	0.512	0.001	[0.510, 0.515]	+1.2	0.008	0.000		87	141,138	376,366
Party	No policy information	0.767	0.012	[0.745, 0.790]	+26.7	0.069	0.000		37	34,145	117,828
Party	Overall	0.718	0.009	[0.701, 0.735]	+21.8	0.077	0.000		78	81,883	385,968
Party	Policy included	0.673	0.008	[0.657, 0.690]	+17.3	0.051	0.000		41	47,738	268,140
Race/ethnicity	No policy information	0.529	0.012	[0.507, 0.552]	+2.9	0.045	0.000		17	22,240	61,767
Race/ethnicity	Overall	0.529	0.007	[0.515, 0.543]	+2.9	0.042	0.000		38	47,475	128,551
Race/ethnicity	Policy included	0.528	0.009	[0.510, 0.546]	+2.8	0.041	0.000		21	25,235	66,784

Note: Estimates cover conjoints from all countries, not US studies only. Full numerical results for every point plotted in Figure 6 of the main text. The dependent variable throughout is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task. $\hat{\mu}$ is the meta-analytic average forced-choice probability of choosing the candidate who matches the respondent on the characteristic over one who does not, so 0.500 is an even split and the premium column reports the same quantity in percentage points. Estimates are computed within study over randomized comparisons and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and Q p the p -value of the heterogeneity test. k counts the study-level estimates pooled; the Overall rows include studies for which policy information was not randomized and so do not appear in either policy row. Standard errors within study are clustered by respondent. Every study-level estimate pooled into a row here appears, one row per estimate, in the estimate-level supplement, Table S3, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

Table Q.23: Estimates underlying Figure 7 (matchups by respondent partisanship and policy information)

Matchup	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Women vs Men	Democrats	No policy information	0.564	0.008	[0.549, 0.580]	0.039	0.000	33	17,934	54,859	
Women vs Men	Republicans	No policy information	0.492	0.003	[0.485, 0.499]	0.006	0.066	32	11,493	34,368	
Women vs Men	Democrats	Policy included	0.522	0.004	[0.514, 0.531]	0.021	0.000	33	21,172	92,287	
Women vs Men	Republicans	Policy included	0.490	0.003	[0.483, 0.497]	0.011	0.018	30	14,062	56,147	
Young vs Old	Democrats	No policy information	0.575	0.013	[0.549, 0.602]	0.045	0.000	17	5,543	9,100	
Young vs Old	Republicans	No policy information	0.533	0.015	[0.505, 0.562]	0.044	0.000	17	3,835	6,400	
Young vs Old	Democrats	Policy included	0.544	0.009	[0.527, 0.561]	0.023	0.000	14	9,466	24,882	
Young vs Old	Republicans	Policy included	0.539	0.005	[0.529, 0.549]	0.005	0.452	14	5,920	16,165	
White vs Black	Democrats	No policy information	0.449	0.010	[0.430, 0.468]	0.036	0.000	24	9,555	15,473	
White vs Black	Republicans	No policy information	0.534	0.010	[0.515, 0.553]	0.028	0.004	22	4,905	8,747	
White vs Black	Democrats	Policy included	0.478	0.009	[0.461, 0.496]	0.041	0.000	31	14,018	28,254	
White vs Black	Republicans	Policy included	0.514	0.007	[0.499, 0.529]	0.026	0.000	28	8,100	16,302	
Christian vs Muslim	Democrats	No policy information	0.572	0.032	[0.509, 0.635]	0.037	0.081	2	531	863	
Christian vs Muslim	Republicans	No policy information	0.754	0.021	[0.713, 0.795]	0.000	0.811	2	395	582	
Christian vs Muslim	Democrats	Policy included	0.574	0.057	[0.462, 0.685]	0.000	1.000	1	143	205	
Christian vs Muslim	Republicans	Policy included	0.643	0.066	[0.513, 0.773]	0.000	1.000	1	109	160	
College Degree vs None	Democrats	No policy information	0.610	0.023	[0.565, 0.655]	0.058	0.000	8	3,580	28,299	
College Degree vs None	Republicans	No policy information	0.594	0.025	[0.544, 0.643]	0.063	0.000	8	2,636	17,236	
College Degree vs None	Democrats	Policy included	0.552	0.024	[0.504, 0.600]	0.044	0.000	4	3,504	62,414	
College Degree vs None	Republicans	Policy included	0.579	0.010	[0.559, 0.598]	0.014	0.044	4	2,068	35,135	
No Experience vs Some	Democrats	No policy information	0.384	0.012	[0.361, 0.408]	0.049	0.000	20	11,744	29,863	
No Experience vs Some	Republicans	No policy information	0.407	0.012	[0.383, 0.431]	0.045	0.000	18	7,088	18,770	
No Experience vs Some	Democrats	Policy included	0.454	0.013	[0.428, 0.480]	0.043	0.000	15	6,786	33,915	
No Experience vs Some	Republicans	Policy included	0.462	0.018	[0.427, 0.497]	0.058	0.000	14	4,676	21,664	
Working Class vs Professional	Democrats	No policy information	0.480	0.017	[0.448, 0.513]	0.024	0.255	7	1,054	2,230	
Working Class vs Professional	Republicans	No policy information	0.469	0.038	[0.395, 0.544]	0.087	0.000	7	643	1,253	
Working Class vs Professional	Democrats	Policy included	0.489	0.008	[0.474, 0.504]	0.010	0.370	6	3,731	9,233	
Working Class vs Professional	Republicans	Policy included	0.506	0.014	[0.479, 0.534]	0.026	0.011	6	2,642	5,837	

Table Q.23: Estimates underlying Figure 7 (matchups by respondent partisanship and policy information) (*continued*)

Matchup	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	$Q\ p$	k	N resp.	N tasks
Republican vs Democrat	Democrats	No policy information	0.183	0.013	[0.157, 0.208]	0.053	0.000	18	10,532	51,688
Republican vs Democrat	Republicans	No policy information	0.804	0.015	[0.775, 0.834]	0.061	0.000	18	6,464	29,123
Republican vs Democrat	Democrats	Policy included	0.334	0.011	[0.313, 0.355]	0.054	0.000	28	19,877	139,134
Republican vs Democrat	Republicans	Policy included	0.671	0.009	[0.653, 0.688]	0.043	0.000	28	14,672	91,934

Note: Estimates are restricted to US studies. Full numerical results for the pooled estimates plotted in Figure 7 of the main text. The figure’s horizontal axis is the Democratic-respondent estimate and its vertical axis the Republican-respondent estimate, so each panel-by-condition cell of the figure is two rows here. The dependent variable throughout is `cand.selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task; $\hat{\mu}$ is the meta-analytic average forced-choice probability that the first-named profile is chosen, so 0.500 is an even split. Estimates are conditional on respondent party identification, computed within study over randomized comparisons and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and $Q\ p$ the p -value of the heterogeneity test. k counts the study-level estimates pooled for that respondent group, which is not in general the panel-level study count the figure annotates: a study contributes to the panel if it yields an estimate for either group. Standard errors within study are clustered by respondent. Estimates for independent respondents are computed but not plotted. Every study-level estimate pooled into a row here appears, one row per estimate, in the estimate-level supplement, Table S4, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

Table Q.24: Estimates underlying Figure 8 (party versus gender tradeoff)

Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Man Democrat	Overall	0.706	0.019	[0.669, 0.743]	0.090	0.000	29	4,513		6,959
Man Democrat	No policy information	0.791	0.019	[0.753, 0.829]	0.051	0.000	14	1,892		2,858
Man Democrat	Policy included	0.646	0.017	[0.612, 0.680]	0.055	0.000	15	2,621		4,101
Man Republican	Overall	0.706	0.019	[0.669, 0.743]	0.087	0.000	29	3,463		5,256
Man Republican	No policy information	0.789	0.022	[0.745, 0.833]	0.064	0.000	14	1,240		1,826
Man Republican	Policy included	0.645	0.011	[0.624, 0.666]	0.012	0.599	15	2,223		3,430
Woman Democrat	Overall	0.679	0.020	[0.641, 0.718]	0.095	0.000	29	5,498		8,489
Woman Democrat	No policy information	0.760	0.028	[0.705, 0.814]	0.090	0.000	14	2,442		3,687
Woman Democrat	Policy included	0.617	0.011	[0.596, 0.638]	0.022	0.159	15	3,056		4,802
Woman Republican	Overall	0.712	0.023	[0.668, 0.757]	0.111	0.000	29	3,336		5,116
Woman Republican	No policy information	0.783	0.035	[0.714, 0.851]	0.116	0.000	14	1,245		1,827
Woman Republican	Policy included	0.652	0.015	[0.622, 0.682]	0.042	0.005	15	2,091		3,289

Note: Estimates are restricted to US studies. Pooled results for the Meta panel of Figure 8 of the main text. The dependent variable is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task. $\hat{\mu}$ is the meta-analytic probability of choosing the profile that matches the respondent's party but not their gender over the profile that matches their gender but not their party, so 0.500 is an even split and values above it favor party. Estimates are computed within study over randomized comparisons, conditional on the respondent's own party and gender, and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and Q p the p -value of the heterogeneity test. k counts the study-level estimates pooled. Standard errors within study are clustered by respondent. The figure also plots each contributing study; those study-level estimates appear, one row per estimate, in the estimate-level supplement, Table S5, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

Table Q.25: Estimates underlying Figure 9 (candidate choice tradeoffs)

Trait	Tradeoff	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Gender	T1	Woman Democrat	Policy included	0.658	0.013	[0.632, 0.685]	0.038	0.002	15		2,942	4,582
Gender	T1	Woman Democrat	No policy information	0.818	0.022	[0.775, 0.860]	0.066	0.000	14		2,404	3,678
Gender	T1	Woman Republican	Policy included	0.659	0.016	[0.628, 0.690]	0.045	0.005	15		2,087	3,238
Gender	T1	Woman Republican	No policy information	0.766	0.023	[0.721, 0.811]	0.064	0.000	14		1,278	1,899
Gender	T1	Man Democrat	Policy included	0.638	0.016	[0.608, 0.669]	0.046	0.000	15		2,645	4,112
Gender	T1	Man Democrat	No policy information	0.767	0.022	[0.725, 0.809]	0.061	0.000	14		1,896	2,893
Gender	T1	Man Republican	Policy included	0.680	0.012	[0.657, 0.703]	0.023	0.155	15		2,260	3,555
Gender	T1	Man Republican	No policy information	0.806	0.028	[0.751, 0.860]	0.085	0.000	14		1,238	1,784
Gender	T2	Woman Democrat	Policy included	0.627	0.018	[0.592, 0.663]	0.058	0.000	15		2,766	4,100
Gender	T2	Woman Democrat	No policy information	0.823	0.021	[0.782, 0.864]	0.064	0.000	14		2,496	3,778
Gender	T2	Woman Republican	Policy included	0.688	0.013	[0.664, 0.713]	0.026	0.113	15		1,891	2,803
Gender	T2	Woman Republican	No policy information	0.812	0.031	[0.751, 0.873]	0.095	0.000	14		1,233	1,879
Gender	T2	Man Democrat	Policy included	0.652	0.016	[0.621, 0.683]	0.050	0.000	15		2,783	4,673
Gender	T2	Man Democrat	No policy information	0.779	0.024	[0.732, 0.827]	0.075	0.000	14		1,921	2,872
Gender	T2	Man Republican	Policy included	0.663	0.013	[0.638, 0.687]	0.030	0.051	15		2,406	4,048
Gender	T2	Man Republican	No policy information	0.783	0.032	[0.721, 0.845]	0.102	0.000	14		1,290	1,941
Gender	T3	Woman Democrat	Policy included	0.659	0.011	[0.639, 0.680]	0.025	0.034	15		3,149	5,290
Gender	T3	Woman Democrat	No policy information	0.810	0.026	[0.759, 0.860]	0.084	0.000	14		2,527	3,866
Gender	T3	Woman Republican	Policy included	0.669	0.010	[0.649, 0.690]	0.018	0.148	15		2,245	3,879
Gender	T3	Woman Republican	No policy information	0.762	0.023	[0.717, 0.807]	0.060	0.002	14		1,262	1,898
Gender	T3	Man Democrat	Policy included	0.660	0.019	[0.623, 0.696]	0.061	0.000	15		2,423	3,694
Gender	T3	Man Democrat	No policy information	0.771	0.031	[0.711, 0.832]	0.098	0.000	14		1,884	2,912
Gender	T3	Man Republican	Policy included	0.651	0.012	[0.627, 0.674]	0.022	0.191	15		2,130	3,129
Gender	T3	Man Republican	No policy information	0.786	0.026	[0.734, 0.837]	0.078	0.000	14		1,241	1,824
Gender	T4	Woman Democrat	Policy included	0.617	0.011	[0.596, 0.638]	0.022	0.159	15		3,056	4,802
Gender	T4	Woman Democrat	No policy information	0.760	0.028	[0.705, 0.814]	0.090	0.000	14		2,442	3,687

Table Q.25: Estimates underlying Figure 9 (candidate choice tradeoffs) (*continued*)

Trait	Tradeoff	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Gender	T4	Woman	Policy included	0.652	0.015	[0.622, 0.682]	0.042	0.005	15		2,091	3,289
		Republican										
Gender	T4	Woman	No policy	0.783	0.035	[0.714, 0.851]	0.116	0.000	14		1,245	1,827
		Republican	information									
Gender	T4	Man Democrat	Policy included	0.646	0.017	[0.612, 0.680]	0.055	0.000	15		2,621	4,101
Gender	T4	Man Democrat	No policy	0.791	0.019	[0.753, 0.829]	0.051	0.000	14		1,892	2,858
			information									
Gender	T4	Man Republican	Policy included	0.645	0.011	[0.624, 0.666]	0.012	0.599	15		2,223	3,430
Gender	T4	Man Republican	No policy	0.789	0.022	[0.745, 0.833]	0.064	0.000	14		1,240	1,826
			information									
Gender	T5	Woman	Policy included	0.527	0.018	[0.492, 0.561]	0.048	0.001	12		2,415	3,674
		Democrat										
Gender	T5	Woman	No policy	0.610	0.022	[0.567, 0.653]	0.059	0.000	12		2,068	3,407
		Democrat	information									
Gender	T5	Woman	Policy included	0.497	0.015	[0.468, 0.527]	0.035	0.021	12		2,386	3,867
		Republican										
Gender	T5	Woman	No policy	0.502	0.012	[0.478, 0.527]	0.000	0.536	11		1,116	1,963
		Republican	information									
Gender	T5	Man Democrat	Policy included	0.489	0.013	[0.464, 0.515]	0.024	0.325	12		2,049	3,207
Gender	T5	Man Democrat	No policy	0.477	0.014	[0.450, 0.503]	0.019	0.180	12		1,507	2,478
			information									
Gender	T5	Man Republican	Policy included	0.541	0.011	[0.519, 0.564]	0.020	0.234	12		2,343	3,831
Gender	T5	Man Republican	No policy	0.500	0.012	[0.476, 0.525]	0.000	0.586	11		1,127	1,972
			information									
Gender	T6	Woman	Policy included	0.551	0.010	[0.532, 0.570]	0.013	0.370	12		2,603	4,345
		Democrat										
Gender	T6	Woman	No policy	0.579	0.013	[0.553, 0.604]	0.026	0.087	11		2,397	4,056
		Democrat	information									
Gender	T6	Woman	Policy included	0.487	0.010	[0.467, 0.508]	0.000	0.907	11		1,767	2,840
		Republican										
Gender	T6	Woman	No policy	0.522	0.011	[0.500, 0.545]	0.000	0.510	11		1,365	2,350
		Republican	information									
Gender	T6	Man Democrat	Policy included	0.486	0.011	[0.466, 0.507]	0.014	0.314	12		2,253	3,828
Gender	T6	Man Democrat	No policy	0.492	0.010	[0.472, 0.511]	0.000	0.250	11		1,852	3,093
			information									
Gender	T6	Man Republican	Policy included	0.519	0.012	[0.495, 0.543]	0.017	0.491	11		1,861	2,916
Gender	T6	Man Republican	No policy	0.522	0.012	[0.499, 0.546]	0.002	0.406	11		1,325	2,181
			information									
Race	T1	Black Democrat	Policy included	0.702	0.028	[0.647, 0.757]	0.063	0.030	11		556	701
Race	T1	Black Democrat	No policy	0.793	0.048	[0.699, 0.888]	0.094	0.000	7		520	741
			information									
Race	T1	Black Republican	Policy included	0.735	0.038	[0.661, 0.810]	0.000	0.688	9		96	135
Race	T1	Black Republican	No policy	0.785	0.118	[0.554, 1.017]	0.168	0.094	5		44	61
			information									
Race	T1	White Democrat	Policy included	0.671	0.023	[0.626, 0.716]	0.067	0.000	11		2,472	3,576

Table Q.25: Estimates underlying Figure 9 (candidate choice tradeoffs) (*continued*)

Trait	Tradeoff	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
Race	T1	White Democrat	No policy information	0.772	0.035	[0.703, 0.840]	0.086	0.000	8	1,066	1,321	
Race	T1	White Republican	Policy included	0.669	0.018	[0.634, 0.704]	0.047	0.000	11	2,528	3,660	
Race	T1	White Republican	No policy information	0.778	0.030	[0.718, 0.837]	0.067	0.000	8	936	1,199	
Race	T2	Black Democrat	Policy included	0.696	0.038	[0.622, 0.770]	0.000	0.691	11	149	165	
Race	T2	Black Democrat	No policy information	0.782	0.071	[0.642, 0.921]	0.139	0.000	7	283	342	
Race	T2	Black Republican	Policy included	0.607	0.146	[0.320, 0.894]	0.258	0.048	6	28	33	
Race	T2	Black Republican	No policy information	0.675	0.102	[0.475, 0.876]	0.111	0.256	4	17	24	
Race	T2	White Democrat	Policy included	0.647	0.011	[0.625, 0.670]	0.020	0.098	11	2,848	6,161	
Race	T2	White Democrat	No policy information	0.805	0.033	[0.740, 0.871]	0.074	0.003	8	531	647	
Race	T2	White Republican	Policy included	0.678	0.009	[0.661, 0.695]	0.011	0.313	11	3,075	6,782	
Race	T2	White Republican	No policy information	0.816	0.037	[0.743, 0.889]	0.081	0.000	8	441	541	
Race	T3	Black Democrat	Policy included	0.668	0.024	[0.621, 0.716]	0.072	0.000	11	1,272	3,423	
Race	T3	Black Democrat	No policy information	0.757	0.039	[0.681, 0.833]	0.087	0.000	8	823	1,552	
Race	T3	Black Republican	Policy included	0.633	0.021	[0.592, 0.675]	0.000	0.749	11	197	631	
Race	T3	Black Republican	No policy information	0.755	0.050	[0.658, 0.853]	0.062	0.560	8	92	156	
Race	T3	White Democrat	Policy included	0.651	0.016	[0.619, 0.682]	0.042	0.001	11	2,286	4,210	
Race	T3	White Democrat	No policy information	0.767	0.030	[0.708, 0.825]	0.073	0.000	8	1,989	3,452	
Race	T3	White Republican	Policy included	0.665	0.012	[0.643, 0.688]	0.023	0.071	11	2,347	4,334	
Race	T3	White Republican	No policy information	0.764	0.029	[0.707, 0.822]	0.070	0.000	8	1,700	3,019	
Race	T4	Black Democrat	Policy included	0.633	0.033	[0.569, 0.697]	0.081	0.003	11	550	690	
Race	T4	Black Democrat	No policy information	0.718	0.036	[0.649, 0.788]	0.059	0.015	8	518	720	
Race	T4	Black Republican	Policy included	0.496	0.044	[0.410, 0.582]	0.002	0.680	8	97	134	
Race	T4	Black Republican	No policy information	0.556	0.124	[0.314, 0.798]	0.213	0.026	5	45	60	
Race	T4	White Democrat	Policy included	0.656	0.012	[0.632, 0.679]	0.023	0.084	11	2,388	3,563	
Race	T4	White Democrat	No policy information	0.776	0.033	[0.711, 0.841]	0.078	0.000	8	1,053	1,349	
Race	T4	White Republican	Policy included	0.628	0.010	[0.609, 0.647]	0.009	0.363	11	2,552	3,714	
Race	T4	White Republican	No policy information	0.778	0.041	[0.697, 0.858]	0.105	0.000	8	906	1,208	
Race	T5	Black Democrat	Policy included	0.614	0.033	[0.549, 0.679]	0.065	0.018	8	431	546	

Table Q.25: Estimates underlying Figure 9 (candidate choice tradeoffs) (*continued*)

Trait	Tradeoff	Respondents	Subsample	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	$Q\ p$	k	N resp.	N tasks
Race	T5	Black Democrat	No policy information	0.715	0.040	[0.636, 0.793]	0.045	0.449	5	228	273
Race	T5	Black Republican	Policy included	0.512	0.044	[0.425, 0.600]	0.032	0.325	8	131	181
Race	T5	Black Republican	No policy information	0.637	0.084	[0.471, 0.802]	0.000	0.900	3	26	35
Race	T5	White Democrat	Policy included	0.513	0.014	[0.485, 0.541]	0.023	0.151	8	1,717	2,446
Race	T5	White Democrat	No policy information	0.453	0.023	[0.408, 0.498]	0.034	0.101	6	1,155	1,362
Race	T5	White Republican	Policy included	0.547	0.017	[0.514, 0.581]	0.039	0.001	9	2,971	4,241
Race	T5	White Republican	No policy information	0.523	0.017	[0.490, 0.557]	0.000	0.534	5	635	822
Race	T6	Black Democrat	Policy included	0.629	0.025	[0.579, 0.678]	0.046	0.107	9	563	758
Race	T6	Black Democrat	No policy information	0.629	0.044	[0.543, 0.715]	0.066	0.125	5	281	367
Race	T6	Black Republican	Policy included	0.616	0.067	[0.485, 0.747]	0.121	0.021	7	85	115
Race	T6	Black Republican	No policy information	0.532	0.080	[0.376, 0.688]	0.002	0.531	5	43	53
Race	T6	White Democrat	Policy included	0.530	0.017	[0.496, 0.564]	0.040	0.002	9	1,987	2,915
Race	T6	White Democrat	No policy information	0.462	0.023	[0.416, 0.507]	0.039	0.051	5	1,148	1,427
Race	T6	White Republican	Policy included	0.542	0.018	[0.507, 0.577]	0.038	0.004	8	2,017	2,894
Race	T6	White Republican	No policy information	0.562	0.037	[0.489, 0.635]	0.073	0.003	5	900	1,147

Note: Estimates are restricted to US studies. Full numerical results for every point plotted in Figure 9 of the main text. The dependent variable is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task; $\hat{\mu}$ is the meta-analytic probability of choosing profile A over profile B in the tradeoff named in the second column, so 0.500 is an even split. The tradeoffs, which appear as facet strips in the figure, are: T1: A: Match on Party, Match on Demographic versus B: Mismatch on Party, Mismatch on Demographic. T2: A: Match on Party, Match on Demographic versus B: Mismatch on Party, Match on Demographic. T3: A: Match on Party, Mismatch on Demographic versus B: Mismatch on Party, Mismatch on Demographic. T4: A: Match on Party, Mismatch on Demographic versus B: Mismatch on Party, Match on Demographic. T5: A: Match on Party, Match on Demographic versus B: Match on Party, Mismatch on Demographic. T6: A: Mismatch on Party, Match on Demographic versus B: Mismatch on Party, Mismatch on Demographic. Estimates are computed within study over randomized comparisons, conditional on the respondent's own party and trait, and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and $Q\ p$ the p -value of the heterogeneity test. k counts the study-level estimates pooled. Standard errors within study are clustered by respondent. Study-level estimates appear, one row per estimate, in the estimate-level supplement, Table S6, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

Table Q.26: Estimates underlying Figure 10 (party-policy tradeoffs)

Tradeoff	Respondents	$\hat{\mu}$	SE	95% CI	$\hat{\tau}$	Q	p	k	N resp.	N tasks
T1	Democrat	0.801	0.020	[0.762, 0.840]	0.081	0.000	26	5,949		11,808
T1	Republican	0.750	0.025	[0.701, 0.800]	0.105	0.000	26	3,890		7,445
T2	Democrat	0.706	0.029	[0.649, 0.763]	0.124	0.000	23	5,783		11,934
T2	Republican	0.734	0.018	[0.700, 0.768]	0.057	0.000	23	3,419		6,839
T3	Democrat	0.729	0.019	[0.692, 0.767]	0.069	0.000	23	5,448		10,873
T3	Republican	0.730	0.022	[0.688, 0.773]	0.081	0.000	22	3,529		7,086
T4	Democrat	0.412	0.031	[0.351, 0.473]	0.134	0.000	23	5,779		11,571
T4	Republican	0.484	0.022	[0.440, 0.528]	0.084	0.000	24	3,854		7,389
T5	Democrat	0.739	0.027	[0.687, 0.791]	0.095	0.000	19	4,671		7,293
T5	Republican	0.674	0.035	[0.605, 0.743]	0.128	0.000	18	3,065		4,716
T6	Democrat	0.752	0.025	[0.703, 0.800]	0.088	0.000	18	4,676		7,196
T6	Republican	0.718	0.034	[0.652, 0.784]	0.117	0.000	17	2,850		4,325

Note: Estimates are restricted to US studies. Full numerical results for every point plotted in Figure 10 of the main text. The dependent variable is `cand_selected`, an indicator for whether the respondent chose the candidate in a binary forced-choice task; $\hat{\mu}$ is the meta-analytic probability of choosing profile A over profile B in the tradeoff named in the first column, so 0.500 is an even split. The tradeoffs, which appear as facet strips in the figure, are: T1: A: Match on Party, Match on Policy versus B: Mismatch on Party, Mismatch on Policy. T2: A: Match on Party, Match on Policy versus B: Mismatch on Party, Match on Policy. T3: A: Match on Party, Mismatch on Policy versus B: Mismatch on Party, Mismatch on Policy. T4: A: Match on Party, Mismatch on Policy versus B: Mismatch on Party, Match on Policy. T5: A: Match on Party, Match on Policy versus B: Match on Party, Mismatch on Policy. T6: A: Mismatch on Party, Match on Policy versus B: Mismatch on Party, Mismatch on Policy. Estimates are computed within study over randomized comparisons, conditional on the respondent's own party identification, and pooled by random-effects meta-analysis; no covariates are adjusted for at any stage, so there are no control coefficients to report. $\hat{\tau}$ is the estimated between-study standard deviation and Q p the p -value of the heterogeneity test. k counts the study-level estimates pooled. Standard errors within study are clustered by respondent. Study-level estimates appear, one row per estimate, in the estimate-level supplement, Table S7, and in machine-readable form at `output/tables/supplement_estimate_level.csv`.

R Use of AI tools

We used Claude Code, Anthropic’s command-line coding agent, between March and October 2026, running successive versions of Anthropic’s Claude models. Our uses of AI go well beyond copyediting, so we describe them here in full. In short, we built the project’s procedures by hand on an initial set of studies; Claude then checked that work and applied the same procedures to a corpus far larger than we could have cleaned ourselves.

Corpus construction. We wrote the cleaning scripts for many of the studies ourselves before Claude was involved. In writing them we settled how each original replication file becomes the standardized respondent, candidate, and task tables, how each study’s attribute labels map onto the standardized categories, and which structural checks a cleaned study must pass before it enters the corpus (two profiles per task, exactly one chosen, every standardized attribute paired with its original label). Claude first checked our hand-written scripts against those rules and against the original replication files, and then followed the same procedure for each study added afterward, writing its cleaning script in the form ours had established. Claude also searched for and screened candidate studies against the eligibility criteria in the main text’s “Meta-reanalysis design” section and recorded the reason for each exclusion in Table B.1; the criteria themselves and every contested inclusion decision are ours. For every study, we reproduced the printed results of its original paper from the replication files and corrected the cleaning script wherever the reproduction disagreed.

Publication audit. We designed the coding scheme for the publication audit: for each study and comparison, whether the original article reports that comparison numerically, only in a figure, or not at all, together with any published value. Claude read the original articles and applied the scheme. These records are the basis of the publication audit in the main text and of the comparison between our re-estimates and the published estimates (Figure E.4).

Analysis and software. The analysis is ours: the estimands, the random-effects specification, weighting, clustering, and the robustness analyses. We wrote the first versions of the core analysis scripts; Claude checked them, revised them as the corpus grew, and wrote the code for the remaining meta-regressions, figures, and tables to our specifications. The same division of labor holds for the R package and the interactive explorer described in Appendix A. We designed both (what each function computes, how the matchup data are organized, and what a user of the explorer can ask), and Claude built the package and

application to our specifications.

Writing. Claude drafted portions of this appendix and of the figure and table notes, proposed revisions to the main text in response to our instructions and to feedback from colleagues, and performed copyediting and style review. We read, edited, and approved all text.

Verification. The large volume studies and estimates offers many opportunities for mistakes; we caught thousands of them through our extensive verification processes. Our re-estimates recover the published estimates closely (Figure E.4), which checks the cleaning scripts against the original authors' own analyses. Every number stated in the main text, this appendix, and the supplement is recomputed from the estimate files by a script that fails if the typeset value differs, and a second script confirms that each value appears in the compiled documents. Every reference was checked against its Crossref record. The full code, including every cleaning script, is in the replication archive, so each step described here can be inspected and rerun.

References

- Abramson, Scott F, Korhan Koçak, and Asya Magazinnik. 2022. “What do we learn about voter preferences from conjoint experiments?” *American Journal of Political Science* 66 (4): 1008–1020.
- Abramson, Scott F, Korhan Kocak, Asya Magazinnik, and Anton Strezhnev. 2026. “Aggregation, Interpretation, and Estimation of Preferences in Conjoint Experiments.” *SocArXiv* preprint, version 3, available at https://doi.org/10.31235/osf.io/xjre9_v3, last accessed August 16, 2026.
- Agresti, Alan, and Brent A. Coull. 1998. “Approximate is Better than ‘Exact’ for Interval Estimation of Binomial Proportions.” *The American Statistician* 52 (2): 119–126.
- Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins, and Teppei Yamamoto. 2021. “Beyond the breaking point? Survey satisficing in conjoint experiments.” *Political Science Research and Methods* 9 (1): 53–71.
- Clayton, Katherine, Yusaku Horiuchi, Aaron R. Kaufman, Gary King, and Mayya Komisarchik. 2026. “Correcting measurement error bias in conjoint survey experiments.” *American Journal of Political Science* p. ajps.70076. Published online July 23, 2026.
- Costa, Mia. 2021. “Ideology, Not Affect: What Americans Want from Political Representation.” *American Journal of Political Science* 65 (2): 342–358.
- Ferrari, Diogo, and Brianna A. Smith. 2025. “Status Threat, Partisanship, and Voters’ Conservative Shift Toward Right-Wing Candidates.” *Journal of Experimental Political Science* 12 (3): 324–338.
- Funck, Amy S, and Katherine T McCabe. 2022. “Partisanship, information, and the conditional effects of scandal on voting decisions.” *Political Behavior* 44 (3): 1389–1409.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge: Cambridge University Press.
- Graham, Matthew H, and Milan W Svobik. 2020. “Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States.” *American Political Science Review* 114 (2): 392–409.
- Leeper, Thomas J, and Joshua Robison. 2020. “More important, but for what exactly? The insignificant role of subjective issue importance in vote decisions.” *Political Behavior* 42 (1): 239–259.
- Lemi, Danielle Casarez. 2021. “Do voters prefer just any descriptive representative? The case of multiracial candidates.” *Perspectives on Politics* 19 (4): 1061–1081.

- Loewen, Peter John, and Ludovic Rheault. 2021. "Voters Punish Politicians with Depression." *British Journal of Political Science* 51 (1): 427–436.
- Mikkelsen, Anna Caroline. 2025. "White Democrats' Growing Support for Black Politicians in the Era of the "Great Awakening"." *American Political Science Review* 119 (4): 1902–1920.
- Miller, David R., and Jeffrey Ziegler. 2024. "Preferential abstention in conjoint experiments." *Research & Politics* 11 (4): 20531680241299329.
- Mummolo, Jonathan, Erik Peterson, and Sean Westwood. 2021. "The limits of partisan loyalty." *Political Behavior* 43 (3): 949–972.
- Pedersen, Helene Helboe. 2026. "Legislator or Representative? Politicians' Tasks According to Voters." *British Journal of Political Science* 56: e4.
- Peterson, Erik. 2017. "The role of the information environment in partisan voting." *The Journal of Politics* 79 (4): 1191–1204.
- Sevi, Semra, Riya Cherian, Jérémie Drouin, Aidee Nunez-Colin, Bradley Wood-MacLean, and Rebecca Zhang. 2025. "Immigration, Abortion, and the Limits of Policy Voting: Evidence from a Candidate Choice Conjoint." Working paper.
- van Oosten, Sanne, Liza Mügge, and Daphne van der Pas. 2024. "Race/Ethnicity in Candidate Experiments: a Meta-Analysis and the Case for Shared Identification." *Acta Politica* 59 (1): 19–41.
- Vevea, Jack L., and Larry V. Hedges. 1995. "A General Linear Model for Estimating Effect Size in the Presence of Publication Bias." *Psychometrika* 60 (3): 419–435.
- Visconti, Giancarlo, and Yang Yang. 2025. "The Limitations of Using Forced Choice in Electoral Conjoint Experiments." Working paper, available at http://www.giancarlovisconti.com/pdfs/Visconti%20and%20Yang.2024.Conjoint_Forced_Choice.pdf, last accessed August 16, 2026.