Policy Legibility and the Limits of Democratic Accountability: Evidence from Voter Beliefs

Saurabh Bhargava
New York University
sbhargav@gmail.com

Daniel Connolly
Princeton University
dan.connolly@princeton.edu

This Version: July 2026

Abstract

Democratic accountability depends on voters' ability to infer how public policies affect their welfare. We study this ability using surveys conducted around the Tax Cuts and Jobs Act of 2017, the 2020 presidential election, and the 2021 American Rescue Plan. Across 11,303 respondents, we compare stated beliefs with respondent-specific statutory benchmarks and common-truth policy facts. We supplement these surveys with candidate-placement and factual-belief measures from the American National Election Studies. We find that welfare-relevant policy consequences are substantially less legible to voters than candidate positions. Beliefs about policy consequences are weakly organized by party, highly dispersed within party, and, for several items, dispersed on a scale comparable to the policy consequence itself. The contrast persists among politically attentive, sophisticated, and affectively engaged voters. An empirically calibrated spatial model maps dispersion of the observed magnitude into electoral incentives and shows that lower legibility substantially flattens the vote-share return to policy moderation. Policy legibility therefore places an important constraint on welfare-based democratic accountability.

JEL Codes. D72, D83, D91

Acknowledgments. We thank Linda Babcock, Larry Bartels, Karna Basu, Lynn Conell-Price, Russell Golman, Kareem Haggag, Daniel Oppenheimer, Silvia Saccardo, participants at the Behavioral Science and Policy and International Behavioral Public Policy conferences, and seminar participants at Carnegie Mellon University for constructive feedback. We also thank Stephanie Rifai, Cassandra Taylor, and several research assistants for excellent project support. The authors retained full editorial discretion and are responsible for any errors.

1 Introduction

Democratic accountability depends, in part, on voters' ability to distinguish the welfare consequences of competing policies. Elections can discipline politicians only if voters can reliably infer how alternative policies affect their own well-being and vote on the basis of this inference. In classic models of spatial competition, candidates are rewarded for moderation because voters can evaluate the policy positions they adopt (Hotelling, 1929; Downs, 1957; Calvert, 1985). These classic models generally assume that voters have their own single-peaked preferences in the policy space and tend to elide the distinction between voters' policy preferences and the welfare consequences of their preferred policies. When welfare consequences are difficult to perceive, however, this distinction becomes important; voters should prefer policies that maximize their welfare, which is difficult to do when one cannot perceive the consequences of policy. Thus, a failure to perceive the consequences of policy weakens an important mechanism through which elections reward or punish politicians, and whether voters accurately perceive policy consequences is therefore a fundamental empirical question. Yet despite a large literature on political knowledge, misinformation, and partisan differences in beliefs, we know surprisingly little about whether voters can accurately infer the welfare consequences of the policies they vote on.

Existing work documents several limits to voters' political information. Voters often know little about politics and rely on parties and elites for informational cues (Delli Carpini and Keeter, 1996; Popkin, 1991; Zaller, 1992). They may misperceive the positions of candidates and parties (Bartels, 1986; Alvarez, 1997; Jessee, 2010), and they frequently hold inaccurate or partisan beliefs about politically relevant facts (Kuklinski et al., 2000; Bartels, 2002; Taber and Lodge, 2006; Bullock et al., 2015; Berinsky, 2018; Jerit and Zhao, 2020). Existing studies also show that citizens often misunderstand taxes, transfers, government benefits, and tax rates in general (Bartels, 2005; Mettler, 2011; Campbell, 2012; Gideon, 2017; Ballard and Gupta, 2018; Stantcheva, 2020*b*). But despite these extensive literatures, research on voter perceptions of policy facts rarely examines beliefs about the effects of policies on respondents'

own household finances.

This omission is unsurprising. Many consequential tax and transfer policies produce effects that depend jointly on household circumstances and complex statutory rules. Determining how such a policy affects a particular household therefore requires combining information about income, filing status, family composition, eligibility, tax brackets, deductions, credits, and phase-out rules. And the same structural complexity that makes these beliefs difficult for researchers to measure may make them difficult for voters to form.

We directly measure these beliefs using original surveys of 11,303 respondents fielded around the Tax Cuts and Jobs Act of 2017 (TCJA), the 2020 presidential election, and the American Rescue Plan Act of 2021 (ARPA). The surveys collect uncommon respondent-level detail about beliefs concerning household tax incidence, stimulus-payment eligibility and magnitude, net government transfers, the distributional incidence and fiscal cost of legislation, and other policy consequences. For household-specific quantities, we construct respondent-specific truth benchmarks using statutory rules and reported household characteristics. For common-truth quantities, we use official administrative and expert sources. Candidate-placement batteries and factual-belief measures from our surveys and the American National Election Studies (ANES) let us compare welfare consequences with other politically relevant objects.

Our central finding is that errors in perceiving policy consequences are large and highly dispersed relative to the policies voters are asked to evaluate. Among respondents with nonmissing belief and benchmark values, the TCJA reduced taxes by an average of $1,642, yet respondents believed their taxes fell by only $173 on average—an average misperception of $1,469. The dispersion in beliefs is equally striking. For the TCJA household-tax item, the residual within-party standard deviation is about $2,502, more than ten times the roughly $225 difference between Democratic and Republican mean errors. In a subsample, we asked respondents how large a tax cut would be required to change their vote. About 30 percent misestimated their own tax change by more than the amount they indicated would

be sufficient. Across nearly every policy quantity, within-party dispersion is larger than the partisan signal available from party identity, and for several items it is comparable to the policy consequence itself.

As a comparative benchmark, we also examine voter beliefs about candidates' placements on ideological and policy scales. Candidate placements are not a one-for-one counterfactual for beliefs about policy consequences, but they concern an object that is central to electoral competition and repeatedly communicated during campaigns. Voters perceive candidate placements on these scales much more consistently: within-party dispersion in placements is smaller than perceived candidate separation, and reversals of the expected candidate ordering are much rarer than errors about whether a policy helps or hurts. This contrast indicates that perceptual difficulty is not uniform across political judgments and that welfare-relevant policy consequences are particularly difficult to discern.

This pattern is not solely a consequence of political disengagement. Politically attentive, sophisticated, and affectively engaged voters exhibit beliefs about policy consequences that are more strongly aligned by party, but are not meaningfully less dispersed within party. In contrast, beliefs about candidate ideologies become more sharply differentiated by party and less dispersed within party with increased political engagement.

Our results could reflect features of our samples or measures rather than genuine differences in what voters perceive. We conduct several robustness checks to address this possibility. First, our original surveys use online convenience samples, whose respondents may hold unusually dispersed beliefs. We therefore reweight each survey to match ANES population margins for demographics, party identification, and household income. The qualitative pattern is unchanged.

Second, the respondent-specific truth benchmarks rely on self-reported household information, which may contain measurement error. We ask how much of the variation in each calculated benchmark would have to be measurement error for voters' underlying beliefs to be fully accurate. The required shares are 92 percent for the TCJA and 75 percent for

the ARPA net transfer, compared with 39 percent for the simpler ARPA stimulus-payment benchmark. Reporting errors would therefore have to be substantially more severe for the more complex policy consequences to explain the weak relationship between beliefs and calculated outcomes. Moreover, we also find substantial belief dispersion for common-truth items, such as the unemployment rate, that do not rely on household-specific inputs.

A third set of concerns involves survey-response artifacts. Extreme entries could inflate dispersion in beliefs, so in our primary specification we winsorize the data while, for robustness, we recompute all estimates without winsorization. Careless or inattentive responding could do the same, and while all surveys included attention checks, we also report results after trimming respondents by completion time. Finally, because our original surveys often elicited continuous or open-ended quantities, whereas the ANES candidate-placement items use bounded categorical scales, we test whether differences in response format drive the contrast. We first coarsen the continuous policy-belief measures into discrete categories and recompute the estimates. We then address the opposite concern—that bounded scales mechanically compress dispersion—by comparing each bounded item's observed dispersion with the dispersion generated by uniform random responding on that item's own scale. None of these adjustments alters the contrast between candidate placements and welfare-relevant policy beliefs, or the robustness of the contrast in high-engagement voters.

We conclude by examining whether policy uncertainty of the observed magnitude matters for electoral competition. We develop an empirically calibrated two-candidate spatial model in which candidates value both electoral support and proximity to their preferred policies, while voters perceive platform consequences with error. The degree of error in the model is disciplined by the empirically-measured dispersion in beliefs, both about candidate ideology and about policy consequences. Belief dispersion in the model set to the empirically-observed candidate-placement dispersion implies a median equilibrium platform distance equal to 56 percent of the total distance between candidate types. Setting it to the level of dispersion observed among policy beliefs, on the other hand, yields final candidate platform distances of

76 percent of the total distance. While our mapping of dispersion from observed welfare beliefs to platform perceptions in the model is not fully structurally estimated, the comparative result is consistent throughout: lower legibility flattens the electoral return to moderation.

We do not argue that policy legibility is the sole determinant of the degree of electoral competition or political polarization. Turnout incentives, activist pressure, primary elections, party brands, affective polarization, and other features of the information environment also clearly shape electoral outcomes. Our contribution is to identify and measure one informational mechanism through which electoral accountability is weakened. When voters cannot infer policy consequences, elections provide weaker incentives for candidates to moderate along that dimension.

The paper contributes to three literatures. First, it provides direct respondent-level evidence on beliefs about the welfare consequences of public policy, extending work on political knowledge, misinformation, policy feedback, and tax-policy perceptions (Delli Carpini and Keeter, 1996; Kuklinski et al., 2000; Mettler, 2011; Campbell, 2012; Jerit and Zhao, 2020; Stantcheva, 2020*a*,*b*). Second, it develops the concept of policy legibility by showing that political objects differ in how readily voters can infer them, complementing work on motivated factual beliefs, partisan economic perceptions, and expressive survey responding (Bartels, 2002; Prior, Sood and Khanna, 2015; Bullock et al., 2015; Berinsky, 2018; Mian, Sufi and Khoshkhou, 2021; Peterson and Iyengar, 2021). Third, it provides an empirical foundation for models in which informational frictions weaken one channel through which elections discipline policy (Diermeier and Li, 2017, 2019; Matějka and Tabellini, 2021; Diermeier and Li, 2023).

2 Data and Policy Background

We use two sources of data. The first is a series of original surveys fielded around three major policy and electoral episodes between December 2017 and March 2021. These surveys elicit beliefs about policy consequences, including household-specific tax and transfer

consequences and common-truth policy facts. The analytic sample retains completed responses from respondents who reported either being registered to vote or having voted in the prior presidential election. The second source is the American National Election Studies (ANES), which provides candidate-placement batteries, unemployment-rate beliefs, and political-knowledge measures in nationally fielded election surveys. Together, these data allow us to compare the legibility of participant-specific welfare consequences of policy, general policy facts, and candidate ideology.

2.1 Original Survey Data

We analyze 11,303 completed responses across the TCJA, the 2020 presidential election, and ARPA. The 2020 election and ARPA samples are restricted to respondents who reported being registered to vote; because the TCJA instruments did not consistently record registration, those samples are restricted to respondents who reported voting in the prior presidential election. All surveys were administered on Qualtrics, with respondents recruited from Amazon Mechanical Turk through the CloudResearch platform. To encourage accuracy, respondents were told that their payment would be doubled if their answer to a pre-selected belief question ranked in the top 10 percent for accuracy.[1] Each survey also collected demographics, party identification, presidential vote where relevant, and feeling-thermometer measures for the major parties or candidates. Full recruitment protocols, elicitation details, and sample composition appear in the Appendix. Because these are online convenience samples rather than probability samples, we also reweight each survey to ANES registered-voter margins in demographics, party identification, and household income. The reweighted estimates preserve the qualitative contrast and rank ordering of the seven headline items; we report the full results in Section 6.[2]

[1]The incentive was common within each survey rather than randomized, so the data do not identify how it affected within-party dispersion. The surveys also do not repeat the same belief item for the same respondent, preventing a test–retest decomposition of persistent belief heterogeneity and transient response error.

[2]Weights are estimated separately by survey wave and capped at 10. We first calibrate to ANES registered-voter margins for age, sex, race and ethnicity, bachelor's-degree attainment, and party identifica-

The surveys elicit two types of policy beliefs. The first type concerns household-specific welfare consequences. For these items, the relevant truth varies across respondents because the effect of policy depends on income, filing status, family composition, and eligibility. We collect data on three such items: the respondent's household tax change under the TCJA, the respondent's own ARPA stimulus payment, and the respondent's net transfer from ARPA. For these items, we construct respondent-specific truth benchmarks using statutory tax and transfer rules and respondents' reported household characteristics. The second type concerns common-truth policy quantities for which the truth is fixed across respondents. These include the distributional incidence of the TCJA, the cost of ARPA as a share of federal spending, the unemployment rate, and the number of new border-wall miles constructed. For these items, we benchmark responses against official administrative data or expert estimates, with sources listed in Table A4.

Household-specific welfare beliefs are the most direct measure of whether voters understand how policy affects their own material circumstances, but they also require the strongest benchmark assumptions. Because benchmark construction is consequential for our estimates, Appendix C documents the statutory calculators, inputs, and sensitivity analyses used to construct respondent-specific truth values. Common-truth items require fewer benchmark assumptions but are less direct measures of household welfare. We use both types of items to characterize the legibility of policy consequences across different informational environments.

2.1.1 The Tax Cuts and Jobs Act of 2017

The Tax Cuts and Jobs Act of 2017 was signed into law on December 22, 2017, after a congressional vote largely along party lines. Its major household-relevant provisions included lower personal income tax rates, a doubled standard deduction, limits on itemization includ-

tion; our preferred specification additionally includes household income. Under the preferred specification, four of the seven headline ratios change by no more than 5 percent. The TCJA household-tax ratio falls from 1.55 to 1.24; the ARPA stimulus-payment and net-transfer ratios rise from 0.68 to 0.77 and from 0.93 to 1.01. The ordering of the seven items is unchanged. Reweighting adjusts for observable differences in sample composition but does not convert the surveys into probability samples or address selection on unobserved characteristics.

ing a $10,000 cap on state and local tax deductions, an expanded Child Tax Credit, and the repeal of personal and dependent exemptions. Contemporaneous estimates suggested that the law would add $1–2 trillion to the federal deficit over the subsequent decade and that roughly two-thirds of the benefit would accrue to the top income quintile (Gale et al., 2018; Tax Policy Center, 2017).

The analytic TCJA sample contains 2,226 respondents across two waves, all of whom reported voting in the prior presidential election: 860 in December 2017 and 1,366 in November 2018. The latter wave coincides with the first major national election after the law took effect. The surveys elicited beliefs about the law's effect on respondents' own tax burden, both as a categorical direction and a dollar magnitude, as well as beliefs about the share of total tax savings accruing to the top and bottom income quintiles. Respondent-specific truths for own-tax-burden questions were calculated from elicited income, filing status, and dependent count using the statutory procedure described in Appendix C.

2.1.2 The 2020 Presidential Election

The 2020 election survey focused on beliefs about policy-relevant facts associated with the two major-party candidates rather than on a single enacted law. We study the change in the U.S. unemployment rate during Donald Trump's first term, the miles of new border wall constructed under his administration, and the effect of Joe Biden's proposed tax plan on a family of four earning an annual household income of $60,000.

The analytic election sample contains 4,861 registered voters across a daily October 2020 panel (n = 1,898) and an Election Day wave (n = 2,963). In addition to the three policy-belief items, the instrument elicited candidate placements on the 1–7 liberal-conservative scale for self, Trump, and Biden; feeling thermometers for both candidates; and a discrete-emotion battery measuring anger, fear, disgust, happiness, hope, and pride toward each candidate. These measures allow us to compare beliefs about policy consequences with beliefs about candidate ideology within the same survey environment.

2.1.3 The American Rescue Plan Act of 2021

The American Rescue Plan Act of 2021 was signed into law on March 11, 2021. It was the fifth major federal COVID-19 response and, at $1.9 trillion, one of the largest U.S. economic stimulus packages. Its household-relevant provisions included direct stimulus payments of $1,400 per person with eligibility phasing out above defined income thresholds, a $300 weekly federal unemployment-insurance supplement, a substantially expanded and partially advanced Child Tax Credit, and a doubled Earned Income Tax Credit for childless households.

The analytic ARPA sample contains 4,216 registered voters surveyed between March 15 and March 27, 2021. The instrument elicited beliefs about respondents' own stimulus payment, their net transfer to or from the government, and the law's total cost as a percent of total U.S. federal government spending in 2020. Respondent-specific stimulus-payment and net-transfer truths were calculated from elicited filing status, income, and dependent count using the procedure described in Appendix C. Respondents with missing required household inputs remain in analyses of the common-truth cost item but are excluded from the respondent-specific items.

2.1.4 Summary of Elicited Beliefs

Across the original surveys, we elicit beliefs for nine distinct policy quantities, with three from the TCJA surveys, three from the 2020 presidential election surveys, and three from the ARPA survey. For three items, own TCJA tax burden, ARPA stimulus payment, and ARPA net transfer, the truth varies across respondents and is computed from elicited income, filing status, and dependent count using the relevant statutory formula. These measures are necessarily limited by the quality of reported household information and by unobserved tax determinants. We therefore report the calculation procedure and sensitivity analyses in Appendix C. For the remaining items, the truth is a fixed value drawn from Tax Policy Center distributional estimates, Bureau of Labor Statistics unemployment figures, Customs

and Border Protection wall-construction records, and Office of Management and Budget historical budget tables. Sources are reported in Table A4.

2.2 American National Election Studies

We supplement the original survey data with the 2016, 2020, and 2024 American National Election Studies (ANES) time-series studies (American National Election Studies, 2019, 2021, 2025). The ANES serves three purposes. First, it provides candidate-placement batteries that allow us to compare the legibility of welfare consequences with the legibility of candidate ideology in a canonical election survey. Second, it includes an objective factual belief item about the U.S. unemployment rate, which allows us to examine whether the dispersion we document for policy beliefs appears in a different sample and elicitation format. Third, it includes civics questions that we use to measure political sophistication.

The candidate-placement batteries ask respondents to locate presidential candidates on common policy scales. Across the three waves, these include liberal-conservative ideology, government services and spending, defense spending, government versus private health insurance, jobs and the standard of living, aid to Black Americans, and environmental regulation versus business, each measured on a 7-point scale, as well as an abortion item. The 2024 wave ($N = 5{,}521$) includes both a 7-point abortion scale and a four-category abortion item; the 2016 ($N = 4{,}270$) and 2020 ($N = 8{,}280$) waves include the four-category abortion item. The 2016 four-category abortion placements were administered post-election and use the ANES post-election weight; the other placement batteries are pre-election and use the corresponding pre-election weight.

The unemployment item asks respondents to identify the pre-election unemployment rate from four options spaced two percentage points apart. The correct rates are 4.9 percent in 2016, 6.9 percent in 2020, and 4.1 percent in 2024. We use this item to assess whether high dispersion in welfare-relevant beliefs also appears outside our original survey instruments.

The civics battery comprises four knowledge questions common to all three waves. They

ask about the smallest category of federal spending, the party holding the House before the election, the party holding the Senate, and the length of a U.S. senator's term. We use these items to measure political sophistication, classifying respondents who answer three or four correctly as high-sophistication and those who answer zero or one correctly as low-sophistication.

Each ANES component plays a distinct role in the analysis. Candidate placements provide the main comparison object for political legibility and supply moments used to calibrate candidate policy motivation in the spatial model. The unemployment item provides an externally fielded benchmark for factual beliefs about a welfare-relevant economic outcome. The civics battery allows us to examine whether belief dispersion is concentrated among low-sophistication respondents or instead persists among more politically knowledgeable voters.

3 Decomposing Belief Errors

We begin by decomposing belief errors into three components: average error, systematic differences across parties, and residual dispersion within party. The purpose of the decomposition is descriptive: it asks whether errors in welfare-relevant beliefs are broadly shared, organized by partisan affiliation, or heterogeneous even among co-partisans. If errors differ systematically across parties, partisan affiliation provides some structure to voters' beliefs. If substantial dispersion remains within party, however, even voters who share the same partisan orientation hold markedly different beliefs about what the policy did.

Let b_i denote respondent i's stated belief about a policy quantity and let τ_i denote the corresponding benchmark truth. For household-specific welfare quantities, τ_i varies across respondents because the policy's effect depends on income, filing status, family composition, and eligibility. For common-truth items, $\tau_i = \tau_q$ is fixed across respondents for item q. We define the belief error as

$$e_i = b_i - \tau_i.$$

For respondent-specific items, using errors rather than raw beliefs removes variation in true policy exposure before characterizing belief dispersion. For common-truth items, the distinction is immaterial because the benchmark truth is constant.

We first illustrate the decomposition using the two major enacted policies for which we elicited household-specific welfare beliefs, the Tax Cuts and Jobs Act of 2017 and the American Rescue Plan Act of 2021. For these items, we relate belief errors to partisan affect using the thermometer difference

$$t_i = \text{GOP thermometer}_i - \text{Democratic thermometer}_i,$$

which ranges from -100 to $+100$. We estimate

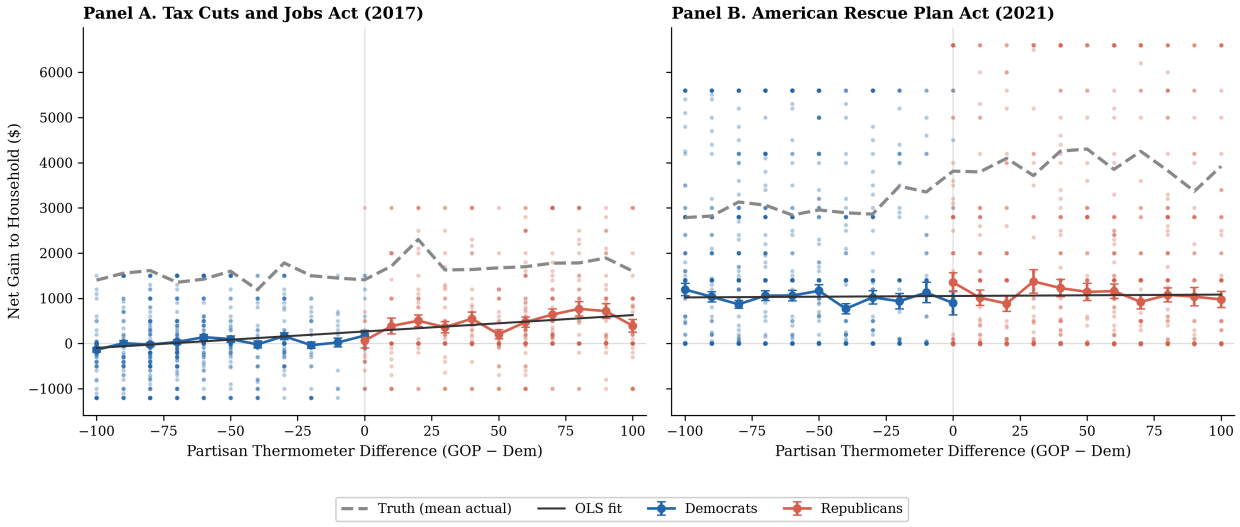
$$e_i = \delta + \beta t_i + \varepsilon_i.$$

Here δ is the expected belief error for a respondent whose partisan thermometer difference is zero, β captures the linear association between partisan affect and belief error, and ε_i is the residual variation not accounted for by that relationship.

Figure 1 plots beliefs about policy consequences over the partisan thermometer difference measure for both TCJA (Panel A) and ARPA (Panel B). Each point is one respondent's stated belief about the policy's net household gain or loss. The figure also reports binned mean beliefs and the fitted relationship between beliefs and partisan affect. For household-specific items, the relevant comparison is between stated beliefs and respondent-specific truths. Accordingly, the visual patterns are purely illustrative; we report the full results of the decomposition below.

The figure highlights three patterns. First, average beliefs fall far below benchmark truths. For TCJA, respondents substantially understate the tax cut they received; for ARPA, they similarly understate the net transfer they received. Second, beliefs move in the expected partisan direction for TCJA. Respondents with warmer feelings toward the GOP report

Figure 1: Voter Beliefs About TCJA and ARPA



Notes. Each scatter point is one respondent. For this visualization, displayed values are additionally winsorized at the 5th and 95th percentiles within each party; binned means and the OLS fit use the same displayed values. Both panels restrict the sample to aligned partisans (Democrats with thermometer difference ≤ 0, Republicans with thermometer difference ≥ 0) for readability. Cleaning-stage processing for the underlying household-dollar beliefs is documented in Appendix B3.

larger gains from a Republican tax law. Third, and most important for our purposes, there is large dispersion in beliefs at nearly every point on the partisan-affect scale. Respondents with similar partisan affect often report very different beliefs about the same policy's household consequences.

We then extend the decomposition to the full set of policy-belief items. To characterize the partisan structure of errors in a comparable way across items, we estimate

$$e_i = \mu + \beta_D(D_i - \bar{D}) + \beta_R(R_i - \bar{R}) + \varepsilon_i,$$

where D_i and R_i indicate Democratic and Republican party identification and $\bar{D}$ and $\bar{R}$ are their means in the estimation sample. Centering the party indicators makes μ the overall mean belief error. The coefficients β_D and β_R remain, respectively, the differences between Democratic and Independent mean errors and between Republican and Independent mean errors, while $\beta_D - \beta_R$ is the Democratic–Republican difference. Table 1 reports the overall and party-specific mean errors, the Democratic–Republican difference, and the residual standard deviation after removing party-specific means.

Table 1 reports the decomposition for all nine policy-belief items. Three features are central. First, average errors are often economically large. Respondents substantially understate the household benefits of both TCJA and ARPA, and they misperceive several common-truth policy quantities by magnitudes that are large relative to the relevant policy scale. The size of the mean error varies across items, so we emphasize the item-specific estimates rather than a single summary characterization.

Second, partisan structure is present but incomplete. For eight of the nine items, the Wald tests reject equality of Democratic and Republican mean errors. The direction is often intuitive, with respondents holding more favorable beliefs about policies or candidates aligned with their party. Yet the partisan component is rarely large enough to organize the full distribution of beliefs.

Table 1: Belief error decomposition.

	(1) Mean belief ($\bar{b}$)	(2) Truth ($\bar{\tau}$)	(3) Mean error ($\hat{\mu}$)	(4) Dem. mean error ($\hat{\mu}_D$)	(5) Rep. mean error ($\hat{\mu}_R$)	(6) Dem.–Rep. gap ($\hat{\mu}_D - \hat{\mu}_R$)	(7) Residual SD($\hat{\varepsilon}$)
Panel A. Tax Cuts and Jobs Act of 2017 (N = 2,170)							
Change in household tax burden	\$173	\$1,642	\$-1,469	\$-1,559	\$-1,334	\$-225	\$2,502
	(32)	(48)	(54)	(78)	(119)	(142)	
Share of cuts to top quintile	51.4%	65.3%	-13.9%	-7.8%	-26.0%	+18.3%***	28.2%
	(0.63)		(0.61)	(0.90)	(1.14)	(1.45)	
Share of cuts to bottom quintile	19.0%	1.0%	+18.0%	+15.8%	+24.4%	-8.6%***	17.4%
	(0.38)		(0.37)	(0.52)	(0.87)	(1.01)	
Panel B. 2020 Presidential Election (N = 4,719)							
Unemployment rate (Sep 2020)	7.0%	7.9%	-0.9%	-0.7%	-1.1%	+0.4%***	2.0%
	(0.03)		(0.03)	(0.04)	(0.05)	(0.07)	
Border wall miles	541	453	+88	-15	+304	-319***	487
	(7)		(7)	(10)	(14)	(18)	
Biden tax plan (\$60K family)	\$-627	\$2,000	\$-2,627	\$-2,014	\$-3,593	\$+1,579***	\$1,676
	(26)		(24)	(35)	(48)	(59)	
Panel C. American Rescue Plan Act of 2021 (N = 4,115)							
Household stimulus payment	\$2,556	\$2,688	\$-133	\$-37	\$-304	\$+266***	\$1,799
	(33)	(32)	(29)	(45)	(60)	(75)	
Household net transfer	\$1,191	\$3,420	\$-2,229	\$-1,842	\$-2,875	\$+1,033***	\$3,196
	(40)	(48)	(52)	(77)	(106)	(131)	
ARPA cost (% fed budget)	32.1%	29.0%	+3.1%	+0.5%	+9.4%	-9.0%***	28.9%
	(0.46)		(0.45)	(0.67)	(0.93)	(1.15)	

Notes. Model. $e_i = \mu + \beta_D(D_i - \bar{D}) + \beta_R(R_i - \bar{R}) + \varepsilon_i$, where $e_i = b_i - \tau_i$, $D_i = 1$ for Democrats, and $R_i = 1$ for Republicans. $\bar{D}$ and $\bar{R}$ are item-specific estimation-sample means. Centering makes $\hat{\mu}$ the overall mean belief error; $\hat{\beta}_D$ and $\hat{\beta}_R$ remain the Democratic–Independent and Republican–Independent differences. Columns (4) and (5) report the implied party-specific mean errors; column (6) reports their difference, $\hat{\beta}_D - \hat{\beta}_R$. $\bar{\tau}$ is the sample mean truth for respondent-specific items and the known scalar truth otherwise (no SE). Standard errors for columns (3)–(6) are HC1 heteroskedasticity robust. Stars in column (6) report two-sided Wald tests of equal Democratic and Republican mean errors ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). For dollar-denominated items, a positive value indicates a net transfer toward the household; a negative value indicates a net transfer away from the household to the government.

Third, residual dispersion is large. Even after accounting for average errors, party-specific deviations, and respondent-specific truths where available, voters within the same party hold widely varying beliefs about the same policy quantities. This residual dispersion is the empirical object that motivates the legibility analysis in the next section.

The household-incidence items illustrate all three features. For the TCJA own-tax-burden item, respondents understate the reduction in their taxes by roughly $1,469 on average. The estimated Democratic–Republican gap is about $225, while the residual standard deviation is $2,502. Within-party dispersion in errors is therefore more than eleven times the partisan gap. For ARPA net transfers, respondents similarly understate their transfers by roughly $2,229 on average. The Democratic–Republican gap is about $1,033, while the residual standard deviation is $3,196. Although the partisan gap is statistically significant, substantial dispersion remains within parties. Party identity moves beliefs in intuitive directions, but it does not make welfare consequences legible.

The TCJA item also allows us to benchmark these errors against respondents' own stated willingness to change their vote. In a subsample of the December 2017 TCJA survey, respondents identified the smallest tax reduction under the other party's plan that would induce them to switch candidates. Among qualifying respondents with a valid response and an observed belief and benchmark, 30.3 percent made an absolute error in perceiving their own tax change that exceeded the dollar threshold assigned to their selected response category.[3] Thus, for a substantial minority, the misperception itself was larger than a policy difference they indicated could change their vote.

More generally, the residual standard deviation is a reduced-form measure of unstructured belief dispersion. It contains variation not absorbed by the other terms in the decomposition, including differences in information environments, numeracy, item interpretation,

[3]Respondents who reported that no amount would induce them to switch, as well as those who selected the open-ended "more than $5,000" category, remain in the denominator and are coded as not meeting the criterion. We conservatively assign $500 to the "less than $500" category and use the listed amount for the remaining finite categories ($N = 860$). Assigning zero rather than $500 to the first category yields 32.6 percent.

policy salience, recall, expressive responding, and errors in constructed truth benchmarks. Although our statutory estimates remove observed variation in true policy exposure for respondent-specific items, unobserved tax and transfer determinants may still contribute to measured dispersion.

Thus, the decomposition establishes two descriptive facts: that partisan identity does not fully characterize variation in welfare-relevant policy beliefs, and large within-party dispersion remains even after removing average errors and party-specific differences. The next section contextualizes the residual dispersion by placing it on interpretable scales, comparing it to partisan belief gaps, candidate-placement beliefs, and the true policy signals voters would need to perceive in order to evaluate policy consequences.

4 Characterizing Policy Legibility in Voter Beliefs

The previous section shows that belief errors about welfare-relevant policy quantities are often large, only partly structured by party, and substantially dispersed within party. We now place that dispersion on interpretable scales. Our goal is to assess whether party identity organizes voters' beliefs and whether the remaining dispersion is small or large relative to the policy consequence voters would need to perceive in order to respond electorally.

We use three complementary statistics. Throughout, σ denotes pooled within-party dispersion, defined formally in Section 4.1. The first two statistics characterize partisan organization rather than accuracy. The noise-to-signal ratio, σ/Δ, characterizes whether voters within the same party disagree more with one another than Democrats and Republicans disagree on average. Party R^2 characterizes how much of the variation in the relevant belief measure is explained by party identity. The third statistic, the swamping ratio σ/D, compares belief dispersion with the true policy signal voters would need to perceive in order to evaluate the policy. Together, these statistics answer three questions: Does party organize beliefs about the policy? Is the partisan signal large relative to within-party dispersion? And is belief dispersion large relative to the policy consequence itself? Because σ/Δ becomes

unstable when the partisan gap is close to zero, our primary statistic for comparisons across political objects is the swamping ratio; σ/Δ and party R^2 characterize the separate question of partisan organization.

4.1 Dispersion Relative to Partisan Signal (σ/Δ)

Let y_i denote the quantity on which we evaluate belief dispersion. For every policy item, we define y_i as the belief error:

$$y_i = b_i - \tau_i.$$

For common-truth items, subtracting the constant benchmark $\tau_i = \tau_q$ does not change dispersion, party gaps, or party R^2, so these statistics would be identical if calculated using stated beliefs. For respondent-specific items, using errors is essential because it removes party differences in true household exposure that could otherwise be mistaken for partisan differences in beliefs.

All party-based statistics in this section are calculated among Democrats and Republicans; Independents are excluded. Let σ denote the pooled within-party standard deviation of y_i, and let Δ denote the absolute difference between Democratic and Republican mean values:

$$\sigma = \sqrt{\frac{(n_D - 1)s_D^2 + (n_R - 1)s_R^2}{n_D + n_R - 2}}, \qquad \Delta = \left|\bar{y}_D - \bar{y}_R\right|.$$

The ratio σ/Δ is a unit-free measure of dispersion relative to partisan separation. Values above one indicate that within-party dispersion exceeds the average difference between parties. In that case, party identity may move beliefs in a predictable direction, but it does not tightly organize what co-partisans believe about the policy quantity.

4.2 Partisan Structure (Party R^2)

Our second statistic is party R^2. For items with one response per respondent, it is the coefficient of determination from

$$y_i = \alpha + \beta R_i + \varepsilon_i,$$

where R_i indicates Republican identification and Democrats are the omitted category.

Candidate-placement items contain separate placements of the Democratic and Republican candidates. For these items, we pool candidate-specific sums of squares rather than stack the two candidates under a common mean or average two candidate-specific R^2 values:

$$R^2_{\text{party}} = 1 - \frac{\sum_{c \in \{D \text{ candidate}, R \text{ candidate}\}} SS_{W,c}}{\sum_{c \in \{D \text{ candidate}, R \text{ candidate}\}} SS_{T,c}},$$

where each candidate's total and within-party sums of squares are centered on that candidate's own overall and party-specific means. Native-survey estimates are unweighted; ANES estimates use the official survey weight for the relevant item. Party R^2 is bounded between zero and one, invariant to the units of the belief measure, and well-defined even when Δ is close to zero. It captures the share of variation explained by the Democratic–Republican distinction.

The two statistics are related but distinct. Party R^2 summarizes the fraction of total variation explained by party, whereas σ/Δ compares within-party dispersion with the distance between party means. Party provides a stronger organizing cue when it explains a meaningful share of variation and when the partisan separation is large relative to within-party dispersion. Neither statistic, however, establishes that the resulting beliefs are accurate.

4.3 Dispersion Relative to the True Policy Signal (σ/D^{policy})

The preceding statistics ask whether party organizes beliefs. Accountability also depends on whether voters can perceive the relevant policy consequence itself. We therefore define the swamping ratio:

$$\frac{\sigma}{D_q^{\text{policy}}}.$$

The numerator is the same pooled within-party dispersion used above. The denominator, D_q^{policy}, is the magnitude of the policy consequence voters would need to perceive for item q relative to the item's relevant reference point or status quo.

For common-truth items, D_q^{policy} is the relevant scalar policy change reported in Table A4, such as the cost of ARPA as a share of federal spending or the number of border-wall miles constructed. For respondent-specific welfare items, the policy consequence varies across households. Let $\mathcal{S}_q$ denote the item-specific sample of Democrats and Republicans with nonmissing beliefs and truth benchmarks, and let $N_q = |\mathcal{S}_q|$. We define

$$D_q^{\text{policy}} = \frac{1}{N_q} \sum_{i \in \mathcal{S}_q} |\tau_i|,$$

the mean absolute respondent-specific policy effect in the same sample used to calculate σ. This aggregation provides an item-level measure of the typical policy consequence faced by respondents in the analysis.

When the swamping ratio is near or above one, within-party belief dispersion is of the same order of magnitude as the policy consequence itself. In this regime, voters may know the partisan direction of a policy but still lack the precision needed to evaluate its welfare consequences.

4.4 Application to Surveyed Policy Items

Table 2 reports these statistics for the nine policy-belief items and the party-thermometer benchmarks from our original surveys. Unlike the three-group decomposition in Table 1, the

Table 2: Measures of policy legibility.

	(1) Overall mean y	(2) Dem mean y	(3) Rep mean y	(4) Partisan gap (Δ)	(5) Within-party SD (σ)	(6) Noise-to-signal (σ/Δ)	(7) Party R^2	(8) Swamping ratio (σ/D)
Panel A. Tax Cuts and Jobs Act of 2017 (N = 1,568)								
Policy beliefs								
Change in household tax burden	\$-1,480	\$-1,559	\$-1,334	\$225	\$2,600	11.54	0.002	1.55
Share of cuts to top quintile	-14.2%	-7.8%	-26.0%	18.3%	27.9%	1.53	0.089	—
Share of cuts to bottom quintile	18.8%	15.8%	24.4%	8.6%	18.0%	2.08	0.050	—
Party thermometer benchmarks								
Dem. Party thermometer (0–100)	58.0	76.9	23.1	53.7	19.4	0.36	0.635	—
Rep. Party thermometer (0–100)	35.1	14.8	72.6	57.8	19.3	0.33	0.672	—
Panel B. 2020 Presidential Election (N = 3,490)								
Policy beliefs								
Unemployment rate (Sep 2020)	-0.9%	-0.7%	-1.1%	0.4%	2.0%	4.72	0.011	0.63
Border wall miles	122	-15	304	319	507	1.59	0.089	1.12
Biden tax plan (\$60K family)	\$-2,693	\$-2,014	\$-3,593	\$1,579	\$1,684	1.07	0.177	0.84
Party thermometer benchmarks								
Dem. Party/candidate thermometer (0–100)	55.6	72.7	32.8	39.9	27.1	0.68	0.347	—
Rep. Party/candidate thermometer (0–100)	40.1	14.2	74.3	60.1	25.3	0.42	0.581	—
Panel C. American Rescue Plan Act of 2021 (N = 2,862)								
Policy beliefs								
Household stimulus payment	\$-151	\$-37	\$-304	\$266	\$1,867	7.01	0.005	0.68
Household net transfer	\$-2,283	\$-1,842	\$-2,875	\$1,033	\$3,229	3.12	0.024	0.93
ARPA cost (% fed budget)	4.2%	0.5%	9.4%	9.0%	29.5%	3.29	0.022	1.02
Party thermometer benchmarks								
Dem. Party/candidate thermometer (0–100)	55.2	76.8	24.9	51.9	21.2	0.41	0.593	—
Rep. Party/candidate thermometer (0–100)	39.5	16.8	71.4	54.6	20.8	0.38	0.625	—

Notes. All calculations use the item-specific sample of Democrats and Republicans with non-missing y_i; Independents are excluded. Panel headers give the total number of Democratic and Republican identifiers in each survey; item-specific sample sizes may be smaller because of missing responses. For policy items, $y_i = b_i - \tau_i$ is the belief error; for thermometer benchmarks, y_i is the reported rating. Columns (1)–(3) report the corresponding overall, Democratic, and Republican means of y_i. Δ is the absolute difference between Democratic and Republican mean values of y_i, and σ is their pooled within-party standard deviation. Party R^2 is the share of variance in y_i explained by Republican identification, with Democrats as the omitted category. σ/D is σ divided by D^{policy}, the relevant policy signal. For common-truth items, D^{policy} is the shared scalar policy change; for respondent-specific items, it is the mean absolute actual household value in the same item-specific sample. Dashes indicate that no external truth benchmark is available. For dollar-denominated items, a positive value indicates a net transfer toward the household and a negative value indicates a net transfer away from the household to the government.

measures here restrict the sample to Democrats and Republicans because they characterize separation between the two major parties. For policy items, columns (1)–(3) report the overall, Democratic, and Republican mean belief errors; for thermometer benchmarks, they report mean ratings. All remaining statistics are calculated from the same item-specific values of y_i. For respondent-specific items, D_q^{policy} is also calculated in that same item-specific sample.

To limit the leverage of extreme open-ended dollar responses, the reported TCJA household-tax and ARPA household-payment and net-transfer beliefs winsorize responses at the 0.5th and 99.5th percentiles. ARPA net-transfer responses are also subject to the documented range screen before winsorization. Bounded items and all truth benchmarks are left unchanged. Appendix Table A5 reports the corresponding unwinsorized estimates, and Appendix B3 documents the processing rules.

Two patterns emerge. First, within-party dispersion exceeds partisan separation for every policy item. The smallest σ/Δ ratio is 1.07 for the Biden tax-plan item; the remaining ratios range from 1.53 to 11.54. Party R^2 reaches 0.177 at most and falls below 0.01 for the TCJA household-tax and ARPA direct-payment items. Thus, even for the item most strongly organized by party, more than 82 percent of individual belief variation remains unexplained by the Democratic–Republican distinction.

Second, these results are not a mechanical feature of how respondents use political survey scales. As a within-survey positive-control benchmark, Table 2 also reports party feeling thermometers. Thermometers measure affective evaluations rather than beliefs about an externally verifiable object, so they do not benchmark perceptual accuracy. They instead show whether party identity coherently organizes responses to another salient political measure. It does: thermometer ratings have σ/Δ well below one and substantially higher party R^2 than the policy-belief items. Respondents who disagree sharply within party about policy consequences are much more similar in their affective evaluations of the parties.

The swamping ratios show that belief dispersion is also large relative to the true scale of

several policies. For the TCJA own-tax-burden item, the mean absolute actual tax change in the analysis sample is approximately $1,675, while the pooled within-party standard deviation of errors is $2,600, yielding a swamping ratio of 1.55. The border-wall ratio is 1.12, and the ARPA-cost ratio is 1.02. The ratios for the unemployment, ARPA direct-payment, Biden tax-plan, and ARPA net-transfer items are 0.63, 0.68, 0.84, and 0.93, respectively. Although below one, these values indicate that belief dispersion remains of the same order of magnitude as the policy signal voters would need to perceive to detect the existence of the policy.

These results establish the main descriptive pattern. Welfare-relevant policy consequences are weakly organized by party, and belief dispersion is often large relative to the scale relevant for evaluation. The thermometer benchmarks show that this dispersion is not mechanically generated by the survey-response setting, although thermometers are affective rather than perceptual measures. The next section therefore turns to candidate placements, which provide a more substantively relevant comparison for political legibility. Placements ask voters to locate external political objects on common ideological and policy scales, allowing us to compare voters' perceptions of those objects with their beliefs about policy consequences.

5 Legibility Across Political Objects

The previous section shows that welfare-relevant policy beliefs are weakly organized by party and highly dispersed within party. We next examine whether that pattern is specific to policy consequences and related factual quantities or instead reflects a general feature of political survey responses. Candidate-placement questions provide a useful comparison because they ask voters to locate the presidential candidates on familiar ideological and policy scales. Although these questions are not directly equivalent to the welfare-belief questions, they allow us to examine whether voters perceive another electorally important political object more clearly.

We draw candidate-placement measures from both our original surveys and the 2016, 2020, and 2024 American National Election Studies (ANES). Across the ANES waves, respondents place the presidential candidates on scales covering liberal–conservative ideology, government services and spending, defense spending, government versus private health insurance, jobs and the standard of living, aid to Black Americans, environmental regulation versus business, and abortion. These measures allow us to compare the legibility of candidate positions with the legibility of welfare-relevant policy quantities.

5.1 Candidate Placements as a Comparison Object

For welfare and policy-belief items, the relevant signal is the true policy magnitude D^{policy}, defined in the preceding section. For candidate placements, no comparable external truth benchmark is available consistently across all items. Our baseline denominator is therefore the separation respondents perceive between the two candidates. Let p_{iDq} and p_{iRq} denote respondent i's placements of the Democratic and Republican candidates on item q. We define

$$D_q^{\text{placement}} = \frac{1}{N_q} \sum_{i \in S_q} |p_{iRq} - p_{iDq}| \,,$$

where S_q is the item-specific sample of Democratic and Republican respondents with valid placements of both candidates. We extend the pooled within-party dispersion measure σ to paired placements by pooling the candidate-specific within-party variances across the four party-by-candidate cells. The candidate-placement swamping ratio is then $\sigma / D_q^{\text{placement}}$.

This baseline ratio is not a measure of placement accuracy: its denominator captures perceived rather than externally validated candidate separation. Instead, computing this ratio admits comparisons beween dispersion in respondents' placements and the candidate separation visible in these same responses. We address the dependence on our choice of denominator below by recalculating the ratios using alternative candidate distances.

Candidate placements are not a one-for-one counterfactual for beliefs about policy consequences. They instead provide a benchmark drawn from another object that is central

to electoral competition. Candidate positions are repeatedly communicated during campaigns and may be easier for voters to observe than household-level policy consequences. If candidate placements exhibit substantially less dispersion relative to the relevant candidate contrast, that would indicate that perceptual difficulty is not uniform across political judgments and that policy consequences are particularly difficult for voters to discern.

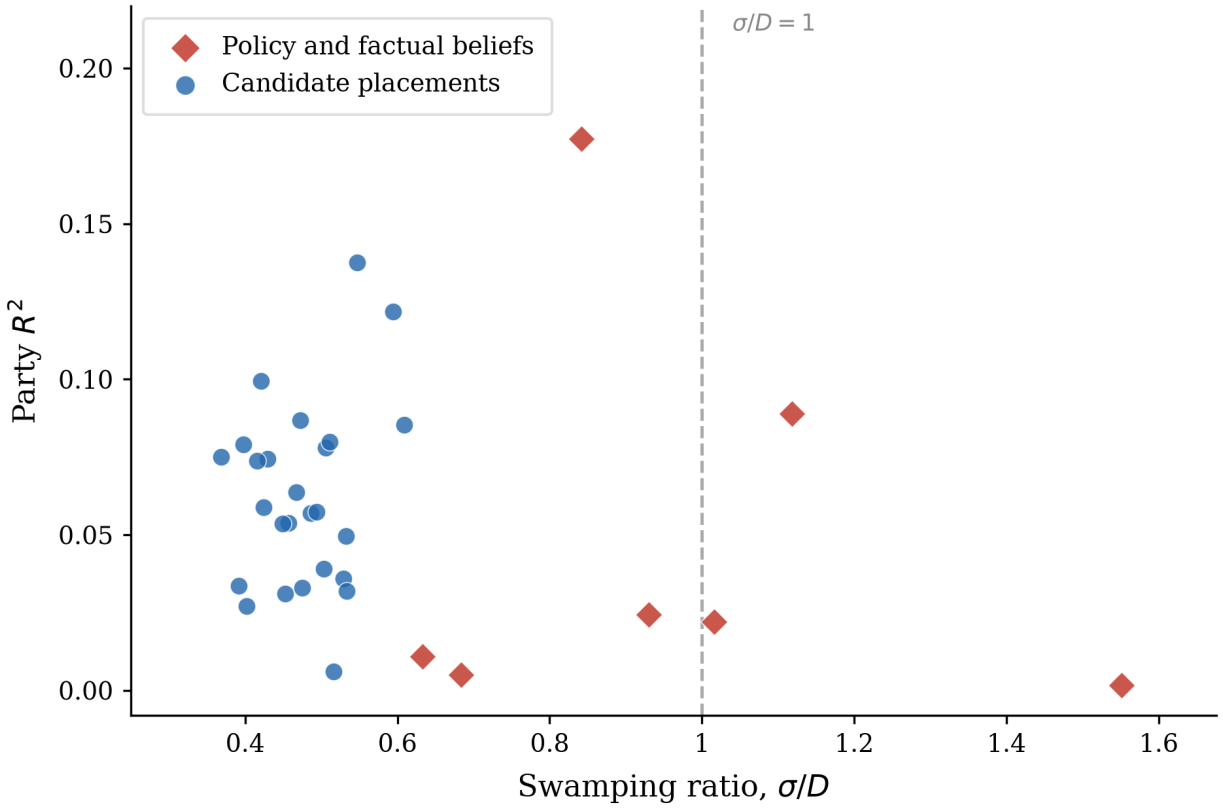
5.2 The Placement–Welfare Asymmetry

Figure 2 places the belief items on two dimensions: the swamping ratio σ/D on the horizontal axis and party R^2 on the vertical axis. For welfare and policy items, $D = D^{\text{policy}}$ is the true policy-induced magnitude. For candidate placements, $D = D^{\text{placement}}$ is the mean absolute perceived separation between the candidates. The comparison therefore gives both ratios the same general interpretation—dispersion relative to the signal voters must distinguish—while recognizing that the underlying signals are defined differently.

The figure shows a clear asymmetry. Across the 26 candidate-placement items, swamping ratios range from 0.37 to 0.61, with a median of 0.47. Thus, within-party placement dispersion is smaller than perceived candidate separation for every placement item. Across the seven policy items, the median ratio is 0.93, and three items have ratios at or above one. The difference between the two class medians is 0.46. A respondent bootstrap that recomputes every item and both class medians yields a 95 percent interval of $[0.41, 0.50]$ for this difference and a two-sided p-value of 0.002. Item-level bootstrap standard errors and intervals appear in Appendix Tables A6 and A7; Appendix Table A8 reports the class-level comparison.

The vertical axis tells a complementary story. The two classes of items overlap considerably in party R^2. Candidate placements are therefore not distinctive primarily because party identification explains much more of their variation. They are distinctive because within-party dispersion is smaller relative to the candidate contrast respondents perceive. Party does not fully organize either class of beliefs, but candidate positions are more legible

Figure 2: Swamping Ratios and Party R^2 Across Belief Items



Notes. Each point represents one belief item. Red diamonds are policy and factual-belief items from the native surveys; blue circles are candidate-placement items from the native surveys and the 2016, 2020, and 2024 ANES. All statistics use item-specific samples of Democratic and Republican respondents; Independents are excluded. The horizontal axis reports the swamping ratio σ/D. For respondent-specific welfare items, σ is the pooled within-party standard deviation of belief errors and D is the mean absolute true household effect in the same item-specific sample. For common-truth items, D is the absolute true policy magnitude. The two TCJA quintile-share items are omitted because they have no defined policy magnitude D. For candidate placements, σ pools candidate-specific within-party dispersion and D is the mean absolute respondent-level perceived candidate separation in the same item-specific sample. The vertical axis reports party R^2 among Democrats and Republicans.

than policy consequences at the scale relevant for political evaluation.

This comparison also weighs against an explanation based solely on generic survey-response noise. The same respondents provide substantially more internally consistent perceptions when evaluating candidate positions. The especially high dispersion in policy beliefs therefore appears connected to the difficulty of translating policy rules, economic conditions, and household circumstances into concrete consequences, although the placement questions are a comparative benchmark rather than a matched counterfactual.

To assess whether the conclusion that candidate placements are more legible than policy consequences depends on the definition of the placement denominator, we report two sensitivity analyses. First, we calculate how often respondents get the direction right for each type of object. Appendix Table A9 compares the two domains on this dimension. For candidate placements, we calculate how often respondents reverse the expected ordering of the Democratic and Republican candidates, where the expected direction follows the response coding and the candidates' party positions. Such reversals are uncommon: 5.3 percent for the native placement item, a median of 9.6 percent across the ANES 7-point items, and 6.1 percent across the ANES 4-point items. The four policy items whose truth benchmark has a directional sign allow for the same test: do participants get the direction of policy consequences correct? On these items, directional accuracy is far lower. Only 33.9 percent of partisans correctly sign their own statutory tax change under the TCJA, with 24.6 percent signing it in the wrong direction and 41.5 percent reporting no change at all. For the Biden tax plan, 55.1 percent sign the consequence backward—believing a tax increase where the truth is a $2,000 cut—while 30.0 percent get the direction right. For the ARPA net transfer, outright sign errors are rare (3.5 percent), but 49.6 percent report a net transfer of zero, leaving 46.9 percent correctly identifying that they are net recipients. The exception is the unemployment rate, where 84.8 percent correctly report that unemployment rose from its January 2017 baseline. Correctly identifying direction is only a minimal test of accuracy, but placements pass it overwhelmingly while welfare consequences largely do not.

Second, we conduct a threshold exercise for placements that does not rely on respondents' perceived candidate separation. We hold the observed within-party dispersion for each placement item fixed and replace its denominator with hypothetical candidate distances ranging from 20 to 50 percent of the item's scale width. For example, on a 1–7 scale with a width of six points, a distance equal to one-third of the scale width corresponds to a two-point separation between the candidates. We then divide the observed placement dispersion by each hypothetical distance.

This exercise does not estimate the candidates' true ideological separation. Instead, it asks how small that separation would have to be for the median placement ratio to equal the policy-item median of 0.93. Appendix Table A10 reports the results. When candidate separation is set to one-third of scale width, the median placement ratio is 0.79. The placement and policy medians become equal when candidate separation is approximately 28.2 percent of scale width. Thus, candidate separation would have to be less than roughly 28 percent of scale width to reverse the placement–policy ordering.

5.3 Heterogeneity by Political Sophistication and Attention

The placement–welfare asymmetry is not confined to politically disengaged respondents. Appendix Table A11 compares low- and high-sophistication ANES respondents. Candidate-placement swamping ratios decline substantially with sophistication: the median σ/D falls from 0.62 to 0.41 in 2016, from 0.60 to 0.36 in 2020, and from 0.51 to 0.33 in 2024. High-sophistication respondents perceive wider candidate separation (median D rises from roughly 2.7–3.1 to 3.4–3.9 scale points) and place candidates with lower within-party dispersion. By contrast, perceptions of the unemployment rate remain highly dispersed relative to the partisan gap: within-party dispersion is several times the Democrat–Republican gap in both sophistication groups in every cycle, and does not improve consistently with sophistication. Sophistication is therefore associated with more sharply differentiated candidate placements

but not with a comparable improvement in perceptions of this factual quantity.[4]

Appendix Table A12 shows that this pattern is, if anything, starker for self-reported political attention in the original surveys. Swamping ratios are essentially unchanged among high-attention respondents: dispersion in welfare-relevant beliefs remains as large relative to the true policy magnitudes as it is for inattentive respondents. What attention changes is partisan structure. Partisan gaps in mean errors roughly double on most common-truth items, and party R^2 rises on every item. Attentive voters absorb partisan signals more strongly without policy consequences becoming more legible relative to their size.

5.4 Heterogeneity by Partisan Affect

Appendix Table A13 examines affective engagement using a bipolar composite of positive emotions toward the respondent's own party and negative emotions toward the opposing party, less the reverse, among Democrats and Republicans in the 2020 survey. Moving from the lowest to the highest affect tertile substantially widens partisan gaps and raises party R^2, while within-party dispersion does not consistently fall and in some cases rises. Swamping ratios remain between 0.6 and 0.9 in the highest tertile. This pattern is consistent with affect operating as a directional heuristic: stronger partisan affect moves beliefs in party-consistent directions without making policy consequences more precisely perceived.

Together, the heterogeneity results reinforce the comparison across political objects. Engagement strengthens candidate differentiation and partisan structure, but it does not consistently reduce within-party dispersion in policy beliefs, and dispersion remains high relative to the size of differences in candidates and policy consequences.

[4]The ANES unemployment outcome is coded in percentage points from each survey's response categories. It has no defined swamping denominator, so the table reports σ and Δ only; across all three waves, within-party dispersion remains large relative to the Democratic–Republican gap in both sophistication groups.

6 Robustness Checks

We next examine whether the placement–welfare asymmetry is an artifact of sample composition, elicitation format, low-effort responding, or measurement error in respondent-specific truth benchmarks.

6.1 Sample Composition

Because our surveys use large convenience samples rather than nationally representative probability samples, observable differences in sample composition could affect estimated within-party dispersion and the resulting ratios. We therefore reweight each survey wave to two sets of registered-voter margins constructed using the ANES 2020 survey weights. The first uses age group, sex, race/ethnicity, bachelor's-degree attainment, and party identification; the preferred specification additionally includes household income in four broad bands. Appendix Table A14 reports both. Under the specification without income, all seven ratios change by at most 4.8 percent. With income, four of seven change by at most 5 percent. The TCJA household-tax ratio falls from 1.55 to 1.24, while the ARPA stimulus-payment and net-transfer ratios rise from 0.68 to 0.77 and from 0.93 to 1.01. The rank ordering of all seven items is unchanged, and the broader placement–policy contrast is preserved. Observable composition therefore affects some magnitudes but does not overturn the main pattern.[5]

6.2 Robustness to Survey Design Choices

Many policy items ask respondents to report cardinal quantities, whereas candidate placements use bounded ordinal scales. The ANES unemployment items provide an independently fielded categorical comparison. Respondents select the unemployment rate from four options,

[5]Weights are estimated separately by survey wave and capped at 10. The income-inclusive specification uses bands below $30,000, $30,000–$59,999, $60,000–$99,999, and $100,000 or more; ANES income refusals and breakoffs are excluded from the income target. All estimates and target margins are restricted to Democrats and Republicans.

a format closer to the candidate-placement elicitations. Empirically, however, the item behaves more like the policy-belief items than the candidate placements. Across 2016, 2020, and 2024, within-party dispersion exceeds the Democratic–Republican gap ($\sigma/\Delta > 1$), and party R^2 is close to zero. High dispersion is therefore not limited to our particular questions or their elicitation format.

One residual concern might be that open-ended continuous response scales, such as the change in one's taxes under TCJA, might be difficult for respondents to interpret and use. To address this possibility, we coarsen the continuous policy-belief items into discrete categories. If the requirement to report exact numerical values were itself responsible for high dispersion, coarsening the responses should make the policy items resemble candidate placements. It does not. For each of the four free-text items from our original surveys, σ/Δ remains above one and party R^2 remains nearly unchanged after coarsening. The results for the other five items, which were elicited using sliders, are similarly unchanged. Appendix Table A15 reports these comparisons.

A parallel concern is that the boundedness of slider and Likert responses could mechanically limit dispersion, potentially giving candidate placements an advantage in the comparison. We therefore compare each bounded item's observed swamping ratio with the ratio that would arise under independent uniform use of its response scale. For each bounded policy or factual item, we calculate the dispersion generated by uniform draws over the full response range and divide it by the same D^{policy} used in the observed ratio. For candidate placements, we draw placements of the two candidates independently from the item's response scale and calculate both the resulting placement dispersion and the expected absolute distance between the candidates.

The observed-to-uniform ratio summarizes how much of this scale-specific random-response dispersion appears in the data. A value of one means that responses are as dispersed, relative to the relevant signal, as independent uniform use of the scale would imply. Lower values indicate that responses are more tightly clustered than that benchmark. The uniform-response

exercise is a reference point for the amount of dispersion permitted by each scale, not a claim that respondents literally guess uniformly.

Table 3: Uniform-response benchmarks for swamping ratios.

	(1) Items	(2) Observed σ/D	(3) Uniform-response σ/D	(4) Observed / uniform-response
Candidate placements				
Native placement [1–7]	1	0.40	0.88	0.46
ANES placements [1–7]	22	0.48	0.88	0.55
ANES placements [1–4]	3	0.48	0.89	0.53
Bounded policy and factual beliefs				
ARPA cost [0–100]	—	1.02	1.00	1.02
Biden tax plan [±$3,000/$5,000]	—	0.84	1.25	0.67
Border wall miles [0–2,000]	—	1.12	1.27	0.88
Unemployment rate [0–10]	—	0.63	0.93	0.68

Notes. Uniform-response benchmarks report the swamping ratio that would arise if responses were independent uniform draws over the elicitation scale. The benchmark uses a continuous uniform draw over the stated response range and divides its standard deviation by the same D^{policy} used in the observed swamping ratio; for the Biden tax item, the benchmark pools the two wave-specific response ranges. The observed/uniform-response ratio in column (4) is independent of D, which cancels. For candidate placements, the observed ratio divides σ by $D^{\text{placement}}$, the mean absolute perceived separation between the candidates; the benchmark draws each candidate independently and uniformly from the ordinal response scale and divides the uniform placement standard deviation by the expected absolute candidate gap under independent random placement. Multi-item placement rows report medians across items.

Table 3 shows that candidate-placement ratios are only 46 to 55 percent of their uniform-response benchmarks. Bounded policy and factual items range from 67 to 102 percent of their corresponding benchmarks. Candidate placements are therefore substantially more concentrated relative to the dispersion permitted by their response scales. The observed asymmetry is not mechanically produced by scale bounds.[6]

Finally, excluding the fastest 5 percent of completions within each wave changes all seven native-survey σ/D ratios by at most 1.1 percent, leaves every classification relative to one unchanged, and preserves the complete ranking.[7]

6.3 Reliability Bounds

A separate concern is that reporting error in household inputs could make the respondent-specific truth benchmarks noisy. We address this concern with an analytical reliability bound.

[6]Open-ended household-dollar items have no natural uniform-response benchmark and are omitted.

[7]The wave-specific thresholds are 240 seconds for TCJA, 246 seconds for the election survey, and 201 seconds for ARPA. Excluding the fastest 10 percent changes ratios by at most 2.5 percent, with the same classifications and ranking.

Under classical measurement error, if respondents knew their true policy consequences perfectly, the slope of stated beliefs on the recorded benchmark would equal the reliability of that benchmark. The ARPA stimulus item provides a conservative reliability floor because its statutory value is a relatively simple function of the same reported household inputs. Attributing the entire attenuation of its belief–truth slope from one to benchmark error yields a reliability benchmark of 0.609.

Explaining the TCJA slope of 0.083 through benchmark error alone would instead require reliability of 0.083—that is, about 92 percent of recorded-benchmark variance would have to be error. If benchmark reliability is at least 0.609, the corrected belief–truth slope is at most 0.14 for TCJA and 0.41 for the ARPA net transfer. Thus, plausible input error can attenuate the estimated relationship, particularly for the net-transfer item, but cannot explain the near-zero TCJA slope. Appendix Table C1 reports the calculation. Common-truth items do not rely on reported household inputs and are unaffected by this concern.

7 Model and Calibration

The empirical results show that candidate ideology is more legible than welfare-relevant policy consequences. We now assess whether dispersion of the observed magnitude can weaken electoral incentives for moderation. The model has two candidates who value both expected electoral support and proximity to their preferred policies, and voters who perceive platform consequences with idiosyncratic error. This approach isolates the role of dispersion in candidate polarization rather than attempting to explain polarization on its own.

7.1 Setup

Candidates L and R have fixed ideological types

$$x_L^\dagger = -0.75, \qquad x_R^\dagger = 0.75,$$

with candidate-type distance $d^{type} = 1.5$. This distance is a normalization, not an empirical estimate. Candidates choose platforms simultaneously, and candidate j maximizes

$$U_j = (1 - \alpha)\, E[V_j(x_j, x_{-j}; \sigma)] - \alpha(x_j - x_j^{\dagger})^2.$$

The first term, $E[V_j(x_j, x_{-j}; \sigma)]$, is expected vote share, and the second, $(x_j - x_j^{\dagger})^2$, is the policy cost of departing from the candidate's bliss point. The expected-vote-share term follows the probabilistic-voting tradition (Hinich, 1977; Lindbeck and Weibull, 1987).[8] The parameter α measures policy motivation relative to office-seeking motivation.

Voter ideal points satisfy $\theta_i \sim U[-1, 1]$. Voters perceive candidate platforms with independent mean-zero error

$$\tilde{x}_{ij} = x_j + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2),$$

and choose the candidate whose platform is perceived to be closer to the voter's ideal point. Candidates know the error distribution. In the baseline model reported here, this independent mean-zero term is the only source of perception error. Appendix Table E2 adds directional and party-linked components calibrated from the Biden tax-plan item.

7.2 How Lower Legibility Weakens Incentives to Moderate

Voter i chooses L when

$$|\theta_i - \tilde{x}_{iL}| < |\theta_i - \tilde{x}_{iR}|.$$

Let $m = (x_L + x_R)/2$ denote the true midpoint between candidate platforms, and let $\tilde{m}_i = (\tilde{x}_{iL} + \tilde{x}_{iR})/2$ denote voter i's perceived midpoint. A voter can mistakenly vote for a candidate whose true platform is farther from her ideal point in one of two ways. Suppose voter i lies to

[8] The Wittman–Calvert tradition also combines policy-motivated candidates with electoral uncertainty (Wittman, 1977, 1990; Calvert, 1985). Its uncertainty is primarily candidate-side and concerns the electorate or its response to platforms. We instead place uncertainty with voters, who imperfectly perceive platform consequences.

the left of the true midpoint, so that $\theta_i < m$. She can reverse the ordering of the candidates, taking R to be the more liberal candidate and voting against her true ranking. Alternatively, she can order the candidates correctly but perceive their midpoint to lie to the left of her ideal point, $\tilde{m}_i < \theta_i$, in which case R appears closer than L.

With independent equal-variance normal errors, the perceived midpoint and perceived candidate ordering are independent. The probability that the ordering flips and the probability that voter i's ideal point lies to the right of her perceived midpoint are

$$B = \Phi\left(\frac{x_L - x_R}{\sigma\sqrt{2}}\right), \qquad A_i = \Phi\left(\frac{\sqrt{2}(\theta_i - m)}{\sigma}\right).$$

The vote probability is therefore

$$P(\text{vote } L \mid \theta_i) = (1 - B)(1 - A_i) + BA_i.$$

Appendix E provides the derivation. When $\sigma = 0$, platforms are perceived correctly and the electoral incentive to approach the median is strongest. As σ rises, vote probabilities move toward one-half and expected vote share becomes less responsive to platform choice. Policy motivation then carries more weight in equilibrium. The model's central comparative static is therefore that lower legibility weakens the electoral return to moderation.

7.3 Model Calibration

Given a level of perceptual noise and a degree of policy motivation, solving the model implies an equilibrium distance between candidate platforms. To calibrate the model, we invert this logic: we take observed placement dispersion from the data and solve for the degree of candidate policy motivation that reproduces the candidate separation voters perceive.

For each ANES issue-year, candidate placements supply two moments. Let $\Delta_{qt}^{candidate}$ denote the mean perceived separation between the Republican and Democratic candidates, after orienting each scale so that positive values correspond to the expected candidate or-

dering. Let $\sigma_{qt}^{placement}$ denote the pooled candidate-specific dispersion in placements. We set the model's relative noise, $\frac{\sigma}{d^{type}}$, equal to the empirical placement ratio $\frac{\sigma_{qt}^{placement}}{\Delta_{qt}^{candidate}}$. The mean perceived candidate separation then supplies the target equilibrium distance.[9]

Let K_{qt} denote the number of scale positions, so that $K_{qt} - 1$ is the full scale range—six points on a seven-point liberal–conservative scale. Dividing perceived separation by $K_{qt} - 1$ expresses it as a share of the scale. Multiplying this share by $d^{type} = 1.5$ then expresses the separation in the model's policy space. The equilibrium-distance target for issue q in year t is

$$D_{qt}^{target} = \left(\frac{\Delta_{qt}^{candidate}}{K_{qt} - 1} \right) d^{type}.$$

This normalization treats the endpoints of the placement scale as the endpoints of the candidates' type space. Because candidate bliss points are not observed directly, Appendix Figure A1 repeats the analysis under alternative candidate-type distances.

The single free parameter α is chosen so that final simulated candidate-platform distance matches this target. Median estimates of α are 0.15 in 2016, 0.19 in 2020, and 0.20 in 2024. Across the 24 issue-year calibrations, the mean and median estimates of α are 0.18. We therefore use $\alpha = 0.175$ as the central specification. Appendix Table A16 reports the issue-year calibrations.

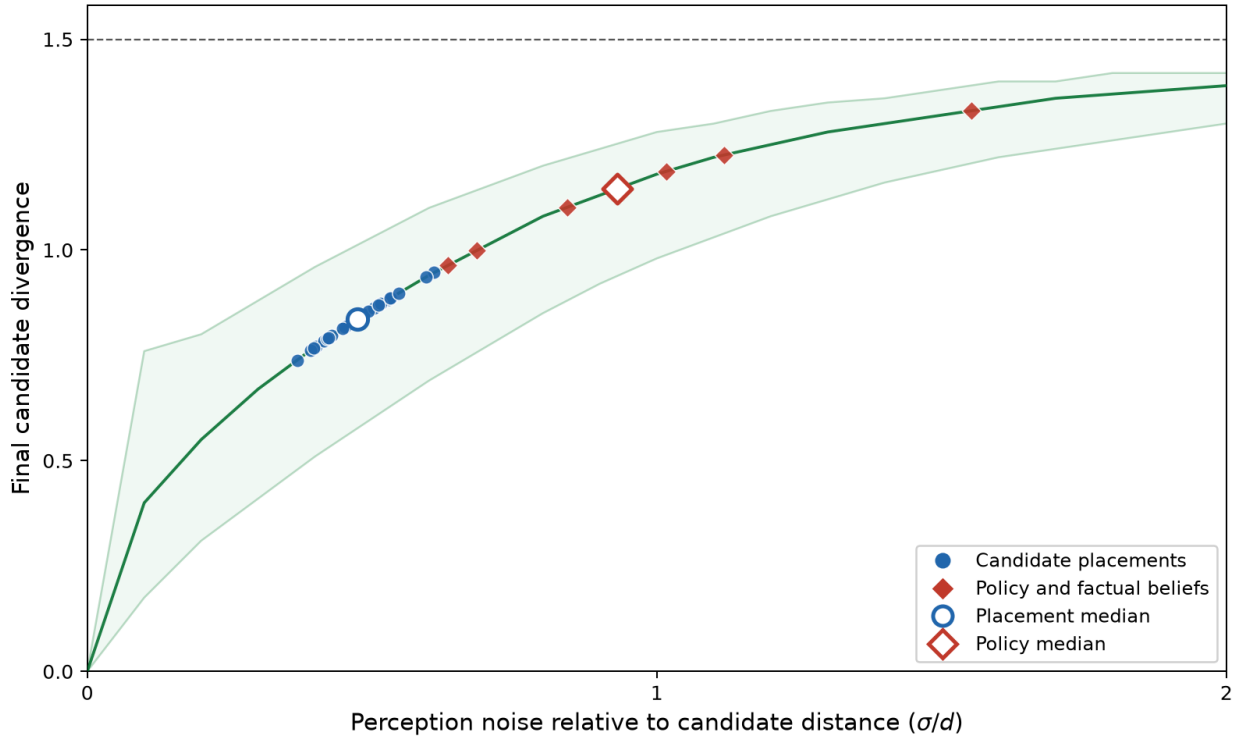
We estimate α from candidate placements rather than policy beliefs because the model's equilibrium condition involves the distance voters perceive between two competing platforms, which placements directly provide. Policy-belief items cannot pin down α in this way. They instead supply noise-to-signal ratios that we take to the model to examine what moderation incentives would look like if platform consequences were perceived with the dispersion with which voters perceive policy consequences.

[9]These calibration moments differ from the placement swamping ratio used in Section 5. The calibration uses pooled placement dispersion across respondents and the mean orientation-adjusted signed candidate separation. The legibility comparison uses within-party dispersion and the mean absolute respondent-level separation. The former supplies unconditional moments for the baseline model and a directional platform-distance target; the latter is designed to compare legibility across political objects net of party.

7.4 Calibrated Simulation Results

We solve the model by iterating best responses on a discrete platform grid. Each parameter configuration uses 100 replications with $N = 1{,}000$ voter ideal points. Appendix E reports the grid, convergence, and tie-breaking rules.

Figure 3: Candidate Platform Divergence by Perceptual Dispersion



Notes. The curve plots median final candidate divergence from the simulation at $\alpha = 0.175$ as a function of perception noise relative to candidate-type distance, σ/d^{type}; the shaded region is the envelope over $\alpha \in \{0.10, 0.175, 0.25\}$. Markers map each belief item's measured swamping ratio σ/D onto the dimensionless model axis: blue circles are candidate-placement items from the native survey and the 2016–2024 ANES, red diamonds are policy and factual-belief items, and large hollow markers are group medians. The dashed horizontal line marks $d^{type} = 1.5$, which bounds equilibrium platform separation in this setup.

Figure 3 shows a monotonic, concave relationship. With little perceptual dispersion, candidates move strongly toward the median. As dispersion rises, equilibrium platforms move toward candidate bliss points because vote shares become less responsive to moderation. The calibrated range of α changes divergence levels but not this relationship.

7.4.1 Mapping the Empirical Measures to the Model

The horizontal axis in Figure 3 reports model perception noise relative to candidate-type distance, σ/d^{type}. To place the empirical items on this curve, we set that dimensionless model ratio equal to each item's empirical swamping ratio, σ/D. For candidate placements, this identifies placement dispersion relative to perceived candidate separation with model noise relative to candidate-type distance. Welfare-relevant policy beliefs have no observed candidate-type distance in dollars, percentages, or policy units. Their locations on the curve are therefore counterfactual: they show how the model behaves when noise is large relative to the relevant signal by the amount observed for policy beliefs, rather than treating a welfare-belief gap as a candidate-platform distance.

To further benchmark the role of belief dispersion in candidate polarization, we compare each bounded item's observed dispersion with the uniform-response benchmark defined in Section 6.2. The observed and uniform values are placed separately on the same model curve; the observed share of uniform-response dispersion is not itself treated as σ/d^{type}.

The electoral consequences of noise in a political object are summarized by how much implied candidate divergence rises when observed dispersion is replaced by uniform responding over the same scale. For candidate placements, the increase is 18 to 23 percentage points of candidate-type distance across the individual items (from 52 to 56 percent observed to 74 to 75 percent under uniform responding); candidates face a correspondingly large electoral penalty for divergence when noise in perceptions is comparable to observed dispersion in candidate placements. For policy items the increase is 3 to 12 percentage points (from 64 to 82 percent observed to 76 to 85 percent uniform); policy-belief dispersion is closer to uniform responding in its equilibrium consequences. This implies a substantially weaker electoral return to moderation. Appendix Table E1 reports the item-level comparisons.

While the exact welfare mapping should be interpreted cautiously, the comparative static is stable. Dispersion of the magnitude observed for policy consequences substantially flattens the electoral return to moderation. Voters may perceive candidate ideology well enough for

it to enter electoral evaluations while remaining unable to evaluate the welfare consequences associated with competing policies at comparable precision.

7.4.2 Robustness to Alternative Assumptions

Appendix Figure A1 shows that the same monotonic result holds under alternative candidate-type distances.

Our baseline model treats perception errors as mean-zero and independent across voters. The survey evidence suggests that errors are neither mean-zero nor independent of party: beliefs about welfare consequences are directionally biased on average and biased differently across parties. Appendix Table E2 therefore re-solves the model under three alternative error structures, with bias moments calibrated from the 2020 election Biden tax-plan item. This item is the closest policy-belief analogue to competing platforms in our data: the first Trump administration enacted its preferred tax policy in 2017, and Biden campaigned on an alternative policy in 2020.

The first alternative error structure adds a common directional shift to all voters' perceived platforms; the second allows the shift to vary across party blocks; and the third combines the two. The calibrated biases are large—the common shift alone is roughly 0.7 times the platform signal—but the comparative static is unchanged. Table E2 reports the results for $\alpha = 0.175$. From $\sigma/d^{type} = 0.5$ through 2.0, equilibrium divergence is weakly increasing at every adjacent grid point under the baseline, common-bias, party-block, and combined specifications; the rank correlation is 1.00 in each case. At $\sigma/d^{type} = 0.9$, the corresponding divergence shares are 0.753, 0.800, 0.813, and 0.840. Directional and party-linked bias shift equilibrium midpoints and raise divergence levels, but the central result holds: increasing residual dispersion weakens candidates' incentives for moderation.

8 Discussion and Conclusion

Voters have difficulty perceiving the welfare consequences of public policy at the scale required for electoral accountability. Across major tax, transfer, and policy episodes, beliefs are weakly organized by party and highly dispersed within party. For several items, dispersion is comparable to the policy consequence itself. Candidate placements clarify that this is not a general inability to perceive political differences: candidate positions are more legible to voters than welfare-relevant policy consequences.

This distinction matters because welfare-based accountability requires more than knowing which party enacted a policy. Voters must infer what the policy did, for whom, and by how much. The calibrated model translates this informational requirement into an electoral consequence. Candidates value both votes and policy, and greater perceptual dispersion flattens the vote-share return to moderation. The model does not claim that legibility is the sole determinant of polarization. Primary elections, activist pressure, turnout incentives, party brands, and affective sorting can sustain divergence even on legible dimensions. The narrower result is that lower legibility weakens one channel through which elections discipline policy.

The findings also connect policy design to political feedback. Policy feedback requires voters to perceive the effects of policy. When voters cannot identify whether or how much a policy benefited them, the path from policy receipt to political reward is weakened at its first stage. The household consequences of changes in taxes and transfers often depend on filing status, dependents, deductions, credits, phase-outs, and other individual circumstances. The difficulty of estimating policy exposure in the present research is therefore not only a measurement challenge; it reflects part of the difficulty voters themselves face in perceiving policy consequences. Policy design may affect not only efficiency and incidence but also whether recipients recognize a policy's effects.

ARPA direct payments provide an illustrative example. Their statutory formula was relatively simple, they were widely discussed in the media, and some survey participants

had received them directly. They are the most legible respondent-specific household policy item in our data. TCJA tax changes, despite having substantial household consequences, are among the least legible. This contrast is suggestive rather than causal, but it is consistent with policy design and the information environment shaping how clearly voters perceive policy consequences.

Political engagement does not eliminate the inference problem voters face. Political sophistication, attention, and affect make beliefs more directionally organized and make candidates appear more differentiated. They do not consistently compress within-party dispersion in beliefs about the welfare consequences of policy. Party identity can therefore provide a directional heuristic without revealing the magnitude of a policy consequence. This distinction explains how beliefs can be partisan and weakly legible at the same time.

Several limitations motivate the paper's robustness checks. The original surveys are not nationally representative probability samples, so we reweight them to ANES registered-voter margins, including income, and compare the results with independently fielded, weighted ANES items. Household truth benchmarks depend on reported characteristics and statutory calculators, so reliability bounds quantify the input error needed to explain the weak belief–truth relationship, while common-truth items avoid household-input error. Unwinsorized estimates, completion-time trims, categorical recoding, and uniform-response benchmarks address sensitivity to survey design and implementation. Alternative candidate-distance denominators test whether the placement comparison depends on treating perceived candidate separation as objective truth. Heterogeneity analyses in the empirical data and directional- and party-bias extensions of the model address partisan-linked and systematic errors. These checks preserve the central contrast: voters perceive candidate separation with less dispersion relative to the relevant signal than they perceive policy consequences.

The broader implication is that policies vary not only in who gains and loses but also in how easily citizens can understand their consequences. Future work should examine whether more legible policies generate stronger feedback, whether information interventions reduce

welfare-belief dispersion, and whether policy domains that differ in legibility also differ in electoral discipline. Elections cannot reliably reward or punish consequences that voters cannot perceive.

References

Alvarez, R. Michael. 1997. *Information and Elections.* Ann Arbor:University of Michigan Press.

American National Election Studies. 2019. "User's Guide and Codebook for the ANES 2016 Time Series Study." September 4, 2019 release.

American National Election Studies. 2021. "User's Guide and Codebook for the ANES 2020 Time Series Study." July 19, 2021 version.

American National Election Studies. 2025. "User's Guide and Codebook for the ANES 2024 Time Series Study."

American Rescue Plan Act of 2021. 2021. *Pub. L. No. 117-2, 135 Stat. 4*, Public Law 117-2.

Ballard, Charles L., and Sanjay Gupta. 2018. "Perceptions and Realities of Average Tax Rates in the Federal Income Tax: Evidence from Michigan." *National Tax Journal*, 71(2): 263–294.

Bartels, Larry M. 1986. "Issue Voting Under Uncertainty: An Empirical Test." *American Journal of Political Science*, 30(4): 709–728.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior*, 24(2): 117–150.

Bartels, Larry M. 2005. "Homer Gets a Tax Cut: Inequality and Public Policy in the American Mind." *Perspectives on Politics*, 3(1).

Berinsky, Adam J. 2018. "Telling the Truth about Believing the Lies? Evidence for the Limited Prevalence of Expressive Survey Responding." *The Journal of Politics*, 80(1): 211–224.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science*, 10(4): 519–578.

Bureau of Labor Statistics. 2020. "The Employment Situation—September 2020." *U.S. Department of Labor news release.*

Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science*, 29(1): 69–95.

Campbell, Andrea Louise. 2012. "Policy Makes Mass Politics." *Annual Review of Political Science*, 15(1): 333–351.

Congressional Budget Office. 2020. "An Update to the Budget Outlook: 2020 to 2030." *Report*.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters.* New Haven:Yale University Press.

Diermeier, Daniel, and Christopher Li. 2017. "Electoral Control with Behavioral Voters." *The Journal of Politics*, 79(3): 890–902.

Diermeier, Daniel, and Christopher Li. 2019. "Partisan Affect and Elite Polarization." *American Political Science Review*, 113(1): 277–281.

Diermeier, Daniel, and Christopher Li. 2023. "The Dynamics of Polarization: Affective Partisanship and Policy Divergence." *British Journal of Political Science*, 53(3): 980–996.

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy*, 65(2): 135–50.

Fabina, Jacob, and Zachary Scherer. 2022. "Voting and Registration in the Election of November 2020." U.S. Census Bureau Current Population Reports P20-585, Washington, DC.

Gale, William G., Hilary Gelfond, Aaron Krupkin, Mark J. Mazur, and Eric J. Toder. 2018. "Effects of the Tax Cuts and Jobs Act: A Preliminary Analysis." *SSRN Electronic Journal*.

Gideon, Michael. 2017. "Do Individuals Perceive Income Tax Rates Correctly?" *Public Finance Review*, 45(1): 97–117.

Hinich, Melvin J. 1977. "Equilibrium in Spatial Voting: The Median Voter Result Is an Artifact." *Journal of Economic Theory*, 16(2): 208–219.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal*, 39(153): 41.

Jerit, Jennifer, and Yangzi Zhao. 2020. "Political Misinformation." *Annual Review of Political Science*, 23(1): 77–94.

Jessee, Stephen A. 2010. "Partisan Bias, Political Information and Spatial Voting in the 2008 Presidential Election." *The Journal of Politics*, 72(2): 327–340.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *The Journal of Politics*, 62(3): 790–816.

Lindbeck, Assar, and Jørgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice*, 52(3): 273–297.

Matějka, Filip, and Guido Tabellini. 2021. "Electoral Competition with Rationally Inattentive Voters." *Journal of the European Economic Association*, 19(3): 1899–1935.

Mermin, Gordon B., Jeff Rohaly, and Jim Nunns. 2020. "An Updated Analysis of Former Vice President Biden's Tax Proposals." Urban–Brookings Tax Policy Center.

Mettler, Suzanne. 2011. *The Submerged State: How Invisible Government Policies Undermine American Democracy.* Chicago:University of Chicago Press.

Mian, Atif, Amir Sufi, and Nasim Khoshkhou. 2021. "Partisan Bias, Economic Expectations, and Household Spending." *The Review of Economics and Statistics*, 1–46.

Office of Management and Budget. 2021. "Historical Tables: Budget of the U.S. Government, Fiscal Year 2022." Executive Office of the President.

Peterson, Erik, and Shanto Iyengar. 2021. "Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading?" *American Journal of Political Science*, 65(1): 133–147.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns.* Chicago:University of Chicago Press.

Prior, Markus, Gaurav Sood, and Kabir Khanna. 2015. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science*, 10(4): 489–518.

Shrider, Emily A., Melissa Kollar, Frances Chen, and Jessica Semega. 2021. "Income and Poverty in the United States: 2020." U.S. Census Bureau Current Population Reports P60-273, Washington, DC.

Stantcheva, Stefanie. 2020*a*. "Understanding Economic Policies: What Do People Know and Learn?" *Working paper*, 156.

Stantcheva, Stefanie. 2020*b*. "Understanding Tax Policy: How Do People Reason?" National Bureau of Economic Research w27699, Cambridge, MA.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science*, 50(3): 755–769.

Tax Cuts and Jobs Act of 2017. 2017. *Pub. L. No. 115-97, 131 Stat. 2054*.

Tax Policy Center. 2017. "Distributional Analysis of the Conference Agreement for the Tax Cuts and Jobs Act." Urban–Brookings Tax Policy Center.

U.S. Customs and Border Protection. 2021. "Border Wall Status Paper."

U.S. Department of the Treasury. 2022. "Fact Sheet: The Impact of the American Rescue Plan after One Year." *Fact sheet*.

Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." *Journal of Economic Theory*, 14(1): 180–189.

Wittman, Donald. 1990. "Spatial Strategies When Candidates Have Policy Preferences." In *Advances in the Spatial Theory of Voting.* , ed. James M. Enelow and Melvin J. Hinich, 66–98. Cambridge:Cambridge University Press.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion.* Cambridge:Cambridge University Press.

Appendix A. Supplemental Tables and Figures

Table A1: Native survey sample demographic summary statistics with U.S. benchmark.

	(1) Full sample	(2) TCJA	(3) Election	(4) ARPA	(5) U.S. reg. voters
Demographics					
Age	39.6***	37.7	38.9	41.3	50.0[a]
	(0.1)	(0.2)	(0.2)	(0.2)	
Male	0.46*	0.46	0.47	0.45	0.47
	(0.00)	(0.01)	(0.01)	(0.01)	
White	0.73***	0.75	0.72	0.75	0.70
	(0.00)	(0.01)	(0.01)	(0.01)	
Employed full-time	0.65	0.73	0.61	0.66	
	(0.00)	(0.01)	(0.01)	(0.01)	
Bachelor's degree	0.64***	0.60	0.64	0.67	0.40
	(0.00)	(0.01)	(0.01)	(0.01)	
Married	0.58***	0.58	—	0.57	0.54
	(0.01)	(0.01)		(0.01)	
Has children	0.41	0.44	—	0.40	
	(0.01)	(0.01)		(0.01)	
Household income ($, self-reported)					
Mean	57,714	60,403	51,147	63,866	
Median	47,500	48,224	37,500	52,500	67,521[c]
25th percentile	25,000	25,000	12,500	27,500	
75th percentile	75,000	75,000	62,500	85,000	
N	11,303	2,226	4,861	4,216	

Notes. Columns pool the sub-waves of each sample (TCJA: 2017 Alabama and 2018 midterm; Election: 2020 daily and Election-Day; ARPA: March 2021). Binary and age rows give the group mean with the standard error beneath in parentheses; income rows summarize self-reported household income (bracket midpoint, dollars). Marital status and children were not collected in the Election wave (—). The U.S. column is the registered-voter composition from the November 2020 CPS (P20-585, Table 2); stars test the pooled (Full sample) estimate against it ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). [a]Approximated from CPS age brackets. Employment has no comparable registered-voter benchmark. [c]Median income of all U.S. households (Census 2020, P60-273); income rows are self-reported household income, and only the median is benchmarked (against all households, not registered voters).

Table A2: Native survey sample political summary statistics with ANES 2020 benchmark.

	(1) Full sample	(2) TCJA	(3) Election	(4) ARPA	(5) ANES 2020
Political participation					
Registered to vote	1.00	—	1.00	1.00	0.89
	(0.00)		(0.00)	(0.00)	
Voted in prev. pres. election	0.87	1.00	0.81	—	0.71
	(0.00)	(0.00)	(0.01)		
Partisanship					
Party identification (share)					
Democrat	0.41***	0.46	0.41	0.40	0.35
Republican	0.29***	0.25	0.31	0.28	0.32
Independent	0.27***	0.27	0.25	0.30	0.30
Feeling thermometers (0–100)					
Democratic Party	51.6***	53.3	51.4	50.9	45.0
	(0.3)	(0.7)	(0.5)	(0.5)	
Republican Party	36.1***	33.3	37.3	36.4	44.7
	(0.3)	(0.7)	(0.5)	(0.5)	
N	11,303	2,226	4,861	4,216	

Notes. Columns pool the sub-waves of each sample (TCJA: 2017 Alabama and 2018 midterm; Election: 2020 daily and Election-Day; ARPA: March 2021). Binary and thermometer rows give the group mean with the standard error beneath in parentheses. Party-identification rows give the share of each sample identifying with the given party (shares need not sum to one because participants may be uncoded). Registration was not consistently collected in TCJA, and prior-election turnout was not collected in ARPA (—); Full-sample participation means therefore use only episodes in which the corresponding item was collected. The ANES column is the weighted American National Election Studies 2020 Time Series benchmark: root party-identification shares, party feeling thermometers, registration status, and recalled 2016 presidential turnout. Stars test the pooled (Full sample) estimate against the benchmark ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$).

Table A3: Tax profile of the TCJA and ARPA native survey samples.

	(1) TCJA	(2) ARPA
Household income ($, self-reported)		
Mean	60,403	63,866
Median	48,224	52,500
25th percentile	25,000	27,500
75th percentile	75,000	85,000
Number of dependents (share)		
0	0.56	0.60
1	0.18	0.18
2	0.16	0.15
3 or more	0.09	0.07
Mean	0.81	0.71
Household size		
Mean number of persons	2.25	2.11
Filing status (share)		
Married filing jointly	0.44	0.45
Married filing separately	0.05	0.03
Single	0.42	0.45
Head of household	0.08	0.07
N	2,226	4,216

Notes. Tax characteristics for the two samples that collected household tax information; TCJA pools the 2017 Alabama and 2018 midterm waves. Household income is the self-reported income bracket midpoint (dollars). Dependents are the number of children in the household. Filing status is tabulated from the raw four-category item, which distinguishes married-filing-separately from single filers (the binary filing indicators used elsewhere collapse the two). ARPA filing-status shares are over the 3,817 respondents who reported a status (399 missing).

Table A4: Survey items, measurement scales, and truth benchmarks.

	(1) Sample size (N)	(2) Measurement scale	(3) Truth type	(4) Truth source
Panel A. Tax Cuts and Jobs Act of 2017 (N = 2,170)				
Change in household tax burden	2,170	Dollars	Individual	Tax Cuts and Jobs Act of 2017 (2017)
Share of cuts to top quintile	2,161	Percent	Common	Tax Policy Center (2017)
Share of cuts to bottom quintile	2,159	Percent	Common	Tax Policy Center (2017)
Panel B. 2020 Presidential Election (N = 4,719)				
Unemployment rate (Sep 2020)	4,719	Percent	Common	Bureau of Labor Statistics (2020)
Border wall miles	4,719	Miles	Common	U.S. Customs and Border Protection (2021)
Biden tax plan ($60K family)	4,719	Dollars	Common	Mermin, Rohaly and Nunns (2020)
Panel C. American Rescue Plan Act of 2021 (N = 4,115)				
Household stimulus payment	3,735	Dollars	Individual	American Rescue Plan Act of 2021 (2021)
Household net transfer	3,735	Dollars	Individual	American Rescue Plan Act of 2021 (2021)
ARPA cost (% fed budget)	4,115	Percent	Common	Office of Management and Budget (2021)

Notes. Individual truths vary across respondents and are computed from elicited household income, filing status, and dependent count using the relevant statutory formula. Common truths are fixed scalars drawn from primary sources. ARPA cost is expressed as a share of FY2020 federal outlays.

Table A5: Robustness to unwinsorized belief responses.

	(1)	(2)	(3)	(4)	(5)
	N	Reported (σ/D)	Unwinsorized (σ/D)	Plausibility screen ($\lvert x\rvert \leq$ \$100,000)	Change (%)
Panel A. Tax Cuts and Jobs Act					
Change in household tax burden	1,568	1.55	2.21	2.21	+42.2
Panel B. 2020 presidential election					
Biden tax plan ($60K family of four)	3,490	0.84	0.84	0.84	0.0
Border wall miles built	3,490	1.12	1.12	1.12	0.0
Unemployment rate	3,490	0.63	0.63	0.63	0.0
Panel C. American Rescue Plan					
Household stimulus payment	2,617	0.68	1.63	1.63	+138.1
Household net transfer	2,617	0.93	1.24	1.24	+33.8
ARPA cost (% of federal spending)	2,862	1.02	1.02	1.02	0.0

Notes. Entries in columns (2)–(4) are swamping ratios σ/D: the pooled within-party standard deviation σ of Democratic and Republican responses (belief errors for respondent-specific items, stated beliefs for common-truth items) relative to the true policy-induced magnitude D, with D held at its canonical value in every column. Column (1) is the number of Democratic and Republican respondents with both a cleaning-stage and a preserved raw response on the item (and, for respondent-specific items, a defined statutory truth); columns (2)–(4) are computed on this common sample. Column (2) uses the cleaning-stage variables employed in the paper. Column (3) substitutes the preserved raw responses. Column (4) restricts the raw responses to $\lvert x\rvert \leq$ \$100,000. Column (5) is the percentage change in σ/D from column (2) to column (3). The 2020 election items and the ARPA cost item are collected on bounded response scales and have no winsorization stage; the ARPA net-transfer cleaning sample imposes the \$100,000 screen before winsorization. † The estimation sample contains 0 raw direct-payment responses above \$100,000; the largest is \$90,000.

Table A6: Candidate placements. Swamping ratio and party R^2.

	(1) N	(2) Swamping ratio (σ/D)	(3) Bootstrap SE	(4) 95% bootstrap interval	(5) Party R^2	(6) Implied divergence ($\alpha = .175$)
Panel A. 2020 Presidential Election Survey						
Liberal–conservative	1,338	0.40	0.012	[0.38, 0.43]	0.027	0.77
Panel B. ANES 2016						
Health insurance coverage	3,612	0.49	0.009	[0.47, 0.50]	0.057	0.85
Aid to Black Americans	3,616	0.49	0.009	[0.48, 0.51]	0.057	0.85
Jobs and standard of living	3,623	0.50	0.009	[0.49, 0.52]	0.039	0.86
Liberal–conservative	3,615	0.51	0.007	[0.49, 0.52]	0.078	0.86
Abortion (4-point scale)	3,123	0.52	0.011	[0.49, 0.54]	0.006	0.87
Government services vs. spending	3,629	0.53	0.010	[0.51, 0.55]	0.036	0.88
Environment vs. business	3,581	0.53	0.009	[0.51, 0.55]	0.032	0.89
Defense spending	3,622	0.61	0.010	[0.59, 0.63]	0.085	0.95
Panel C. ANES 2020						
Health insurance coverage	7,223	0.42	0.007	[0.40, 0.43]	0.074	0.78
Liberal–conservative	7,230	0.43	0.007	[0.42, 0.44]	0.074	0.80
Environment vs. business	7,201	0.45	0.008	[0.44, 0.47]	0.031	0.82
Jobs and standard of living	7,222	0.46	0.007	[0.44, 0.47]	0.054	0.82
Aid to Black Americans	7,212	0.47	0.008	[0.46, 0.49]	0.087	0.84
Abortion (4-point scale)	7,132	0.48	0.008	[0.46, 0.49]	0.033	0.84
Government services vs. spending	7,228	0.53	0.008	[0.52, 0.55]	0.050	0.89
Defense spending	7,217	0.55	0.008	[0.53, 0.56]	0.137	0.90
Panel D. ANES 2024						
Liberal–conservative	4,684	0.37	0.008	[0.35, 0.38]	0.075	0.74
Abortion (7-point scale)	4,674	0.39	0.009	[0.37, 0.41]	0.034	0.76
Abortion (4-point scale)	4,661	0.40	0.009	[0.38, 0.41]	0.079	0.77
Aid to Black Americans	4,645	0.42	0.009	[0.40, 0.44]	0.099	0.79
Environment vs. business	4,630	0.42	0.009	[0.41, 0.44]	0.059	0.79
Jobs and standard of living	4,648	0.45	0.009	[0.43, 0.47]	0.054	0.81
Health insurance coverage	4,651	0.47	0.009	[0.45, 0.49]	0.064	0.83
Government services vs. spending	4,679	0.51	0.010	[0.49, 0.53]	0.080	0.87
Defense spending	4,645	0.59	0.010	[0.58, 0.61]	0.122	0.94

Notes. Native survey: 2020 election wave (MTurk). ANES: 2016, 2020, and 2024 time-series. Candidate-placement R^2 pools sums of squares across the Democratic- and Republican-candidate placement variables. σ/D is a swamping ratio: the pooled within-party standard deviation σ relative to the magnitude of the relevant signal D, not a partisan gap. D is $D^{\text{placement}}$, the mean absolute perceived separation between the Democratic and Republican candidates. Party R^2 is the share of item-level variance explained by Democratic versus Republican identification; Independents are excluded and ANES leaners are counted with parties. ANES statistics use survey weights for means, standard deviations, and R^2; native-survey statistics are unweighted. Implied divergence is the median final platform distance from the simulation at each item's σ/D, evaluated at $\alpha = .175$; the bliss-point bound is $d^{type} = 1.5$. Bootstrap SEs and percentile intervals are from 1,000 respondent resamples, with σ/D recomputed in each resample. Items are sorted by σ/D ascending within each sub-section.

Table A7: Policy and factual beliefs. Swamping ratio and party R^2.

	(1) N	(2) Swamping ratio (σ/D)	(3) Bootstrap SE	(4) 95% bootstrap interval	(5) Party R^2	(6) Implied divergence ($\alpha = .175$)
Panel A. Tax Cuts and Jobs Act						
Change in household tax burden	1,568	1.55	0.084	[1.39, 1.71]	0.002	1.33
Panel B. 2020 presidential election						
Biden tax plan ($60K family of four)	3,490	0.84	0.012	[0.82, 0.87]	0.177	1.10
Border wall miles built	3,490	1.12	0.015	[1.09, 1.15]	0.089	1.23
Unemployment rate	3,490	0.63	0.008	[0.62, 0.65]	0.011	0.96
Panel C. American Rescue Plan						
Household stimulus payment	2,617	0.68	0.020	[0.65, 0.72]	0.005	1.00
Household net transfer	2,617	0.93	0.022	[0.89, 0.98]	0.024	1.14
ARPA cost (% of federal spending)	2,862	1.02	0.013	[0.99, 1.04]	0.022	1.19

Notes. Panels A–C native MTurk waves; items within each wave follow their natural order rather than σ/D. σ/D is a swamping ratio: the pooled within-party standard deviation σ relative to the magnitude of the relevant signal D, not a partisan gap. D is D^{policy}, the true policy-induced magnitude (a scalar truth for common-truth items, the mean absolute respondent-specific effect for welfare items). For respondent-specific welfare items, σ is computed from belief errors; for common-truth items, σ is computed on stated beliefs. Party R^2 is the share of item-level variance explained by Democratic versus Republican identification; Independents are excluded. All statistics are unweighted. Implied divergence is the median final platform distance from the simulation at each item's σ/D, evaluated at $\alpha = .175$; the bliss-point bound is $d^{type} = 1.5$. Bootstrap SEs and percentile intervals are from 1,000 respondent resamples, with σ/D recomputed in each resample.

Table A8: Bootstrap comparison of policy and candidate-placement swamping ratios.

	(1) Items	(2) Estimate	(3) Bootstrap SE	(4) 95% bootstrap interval	(5) p-value
Policy-item median	7	0.930	0.022	[0.889, 0.976]	—
Candidate-placement median	26	0.474	0.005	[0.464, 0.485]	—
Difference: policy minus placement	—	0.456	0.023	[0.414, 0.503]	0.002

Notes. The first two rows report medians across the seven policy items and 26 candidate-placement items in Figure 2. The final row is the policy median minus the placement median. Bootstrap statistics use 1,000 respondent resamples. Respondents are resampled once within each native-survey or ANES wave in every replicate, preserving cross-item dependence within wave. Each σ/D and both class medians are then recomputed. The reported p-value is a finite-replicate, two-sided bootstrap test of a zero class-median difference.

Table A9: Directional errors in candidate placements and policy beliefs.

	(1) Items	(2) Reversed order or wrong direction (%)	(3) Tied or no change (%)	(4) Correct response (%)
Panel A. Candidate placements				
Native 1–7 placement	1	5.3	7.0	87.7
ANES 1–7 placements	22	9.6	11.7	79.2
ANES 1–4 placements	3	6.1	12.1	80.7
Panel B. Policy and factual beliefs				
Change in household tax burden	—	24.6	41.5	33.9
Biden tax plan ($60K family)	—	55.1	14.9	30.0
Unemployment rate (Sep 2020)	—	13.1	2.1	84.8
Household net transfer	—	3.5	49.6	46.9

Notes. For candidate placements, a correct response places the Republican candidate on the side of the Democratic candidate implied by the question coding, a reversed response places the Republican candidate on the opposite side, and ties place both candidates at the same point; shares are computed over paired placements among Democratic and Republican respondents, and placement rows report medians across items within each scale group. For policy beliefs, a correct response is a stated belief with the same sign as the truth benchmark, wrong direction is the opposite sign, and no change is a stated belief of exactly zero. For respondent-specific items (household tax burden, household net transfer), the benchmark is the respondent's own statutory value and respondents with a zero statutory value are excluded. The unemployment item is the change from the January 2017 baseline stated in the instrument. Items whose truth benchmark has no sign (household stimulus payment, border wall miles, ARPA cost, quintile shares) are omitted. ANES placements use survey weights; native calculations are unweighted.

Table A10: Candidate-placement denominator sensitivity.

Panel A. Perceived candidate separation

Sample and scale	Items	Median perceived gap (% of scale width)	Median observed σ/D
Native 1–7 placement	1	55.0	0.402
ANES 1–7 placements	22	56.1	0.479
ANES 1–4 placements	3	62.3	0.475

Panel B. Externally specified candidate separation

External separation (% of scale width)	Median placement σ/D	Placement minus policy median
20.0	1.313	+0.383
25.0	1.050	+0.120
33.3	0.788	-0.142
50.0	0.525	-0.405
Break-even separation: 28.2	0.930	0.000

Notes. Scale width is the distance between the response-scale endpoints. Panel B replaces each item's perceived-separation denominator with a common externally specified share of that item's scale width while retaining its observed within-party placement dispersion. The policy-item median is 0.930. The break-even row is the external separation at which the median placement ratio equals that policy median. ANES calculations use survey weights; the native placement is unweighted.

Table A11: Legibility measures by political sophistication (ANES 2016, 2020, 2024).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Within-party SD (σ)		Partisan gap (Δ)		Denominator (D)		Swamping (σ/D)		Party R^2	
	Low	High	Low	High	Low	High	Low	High	Low	High
Panel A. 2016 candidate placements (N = 953 low; 1,372 high)										
Liberal–conservative	1.80	1.41	1.48	2.56	3.10	3.33	0.58	0.42	0.057	0.111
Government services vs. spending	1.74	1.33	1.57	2.82	2.70	3.22	0.64	0.41	0.010	0.106
Defense spending	1.63	1.40	1.29	1.89	2.37	2.59	0.69	0.54	0.065	0.139
Health insurance coverage	1.76	1.43	1.90	3.27	2.92	3.76	0.60	0.38	0.035	0.111
Jobs and standard of living	1.73	1.36	1.70	3.06	2.75	3.51	0.63	0.39	0.009	0.098
Aid to Black Americans	1.65	1.40	2.25	3.04	2.80	3.39	0.59	0.41	0.029	0.102
Environment vs. business	1.75	1.44	1.39	3.10	2.51	3.56	0.70	0.40	0.016	0.073
Abortion	1.01	0.79	1.24	1.74	1.71	1.86	0.59	0.43	0.004	0.023
Median	1.74	1.40	1.53	2.93	2.72	3.36	0.62	0.41	0.023	0.104
Panel B. 2020 candidate placements (N = 1,754 low; 3,219 high)										
Liberal–conservative	1.74	1.17	2.01	3.18	3.29	3.48	0.53	0.34	0.059	0.126
Government services vs. spending	1.89	1.35	1.60	3.09	2.83	3.45	0.67	0.39	0.030	0.126
Defense spending	1.71	1.36	1.42	2.37	2.58	2.93	0.66	0.47	0.100	0.174
Health insurance coverage	1.77	1.23	2.39	3.69	3.22	3.95	0.55	0.31	0.027	0.155
Jobs and standard of living	1.81	1.27	2.24	3.42	3.08	3.72	0.59	0.34	0.011	0.134
Aid to Black Americans	1.77	1.35	2.49	3.34	3.02	3.67	0.59	0.37	0.026	0.174
Environment vs. business	1.85	1.29	2.03	3.79	2.96	4.08	0.62	0.31	0.009	0.077
Abortion	1.02	0.74	1.13	1.82	1.68	2.00	0.61	0.37	0.015	0.067
Median	1.77	1.28	2.02	3.26	2.99	3.58	0.60	0.36	0.026	0.130
Panel C. 2024 candidate placements (N = 1,651 low; 1,526 high)										
Liberal–conservative	1.62	1.25	2.97	3.53	3.92	4.00	0.41	0.31	0.054	0.108
Government services vs. spending	1.78	1.32	1.77	3.32	2.93	3.73	0.61	0.35	0.059	0.127
Defense spending	1.63	1.58	1.63	1.48	2.69	2.73	0.60	0.58	0.107	0.163
Health insurance coverage	1.67	1.34	2.40	3.47	3.09	3.81	0.54	0.35	0.031	0.146
Jobs and standard of living	1.61	1.27	2.38	3.57	3.13	3.86	0.52	0.33	0.042	0.105
Aid to Black Americans	1.53	1.26	2.77	3.66	3.24	3.90	0.47	0.32	0.076	0.158
Environment vs. business	1.61	1.19	2.48	3.96	3.16	4.20	0.51	0.28	0.059	0.094
Abortion*	1.64	1.16	2.97	3.86	3.62	4.08	0.45	0.28	0.023	0.088
Median	1.62	1.27	2.44	3.55	3.14	3.88	0.51	0.33	0.057	0.117
Panel D. ANES unemployment-rate knowledge item										
2016	1.65	1.25	0.29	0.21	—	—	—	—	0.007	0.018
2020	2.10	1.92	0.73	0.72	—	—	—	—	0.024	0.032
2024	1.69	1.25	0.48	0.34	—	—	—	—	0.026	0.020

Notes. Sophistication is the number of correct answers on a four-item ANES civic knowledge battery (senator term length, least-funded federal program, pre-election House and Senate majority party). Low (0–1 correct) and high (3–4 correct) respondents are compared, using thresholds fixed across cycles; respondents with exactly two correct, or who did not answer all four items, are omitted. For candidate placements, σ is the pooled within-party standard deviation of individual candidate placements, Δ the gap between the two candidates' weighted mean placements within the group, D the weighted mean absolute perceived separation between the Democratic and Republican candidates among Democratic and Republican respondents, σ/D the swamping ratio, and party R^2 pools sums of squares across both candidate placements. For the unemployment item, responses are coded in percentage points, Δ is the absolute Democrat–Republican gap in mean responses, and no swamping denominator is defined. All statistics use ANES sampling weights. *The 2024 abortion entry uses the 7-point placement item; 2016 and 2020 use the 4-point item.

Table A12: Legibility measures by self-reported attention to national politics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Within-party SD (σ)		Partisan gap (Δ)		Denominator (D)		Swamping (σ/D)		Party R^2	
	Low	High	Low	High	Low	High	Low	High	Low	High
Panel A. Tax Cuts and Jobs Act of 2017 (N = 727 low; 1,499 high)										
Change in household tax burden	2,345	2,704	220	213	1,506	1,749	1.56	1.55	0.002	0.001
Share of cuts to top quintile	27.81	27.52	10.95	20.73	—	—	—	—	0.037	0.110
Share of cuts to bottom quintile	17.54	18.07	5.25	9.98	—	—	—	—	0.021	0.062
Panel B. 2020 Presidential Election (N = 2,050 low; 2,805 high)										
Unemployment rate (Sep 2020)	2.03	1.90	0.15	0.56	3.10	3.10	0.65	0.61	0.001	0.021
Border wall miles built	493	512	224	381	453	453	1.09	1.13	0.049	0.119
Biden tax plan ($60K family of four)	1,610	1,706	1,116	1,879	2,000	2,000	0.80	0.85	0.106	0.228
Panel C. American Rescue Plan Act of 2021 (N = 1,491 low; 2,725 high)										
Household stimulus payment	1,625	1,975	246	279	2,839	2,678	0.57	0.74	0.006	0.005
Household net transfer	3,221	3,219	872	1,048	3,711	3,356	0.87	0.96	0.018	0.025
ARPA cost (% of federal spending)	29.43	29.48	5.80	10.67	29.00	29.00	1.01	1.02	0.010	0.030

Notes. All statistics are computed on errors (stated belief minus the truth benchmark; for respondent-specific items, each household's own statutory value). σ is the pooled within-party standard deviation of errors, Δ the absolute Democrat–Republican gap in mean errors, D the swamping denominator (the within-group mean absolute statutory truth for respondent-specific items; the fixed policy magnitude for common-truth items), σ/D the swamping ratio, and party R^2 the share of error variance explained by party among Democrats and Republicans, as in Table 2. Quintile-share items are bounded shares with no defined swamping denominator. Attention splits are wave-specific. TCJA: high attention is following politics "all the time" or "most of the time". 2020 Election: high attention is 8 or above on the 1–10 election-attention scale, pooling the daily-tracking and Election-Day waves. ARPA: high attention is reporting attention to national politics "most of the time" or "all of the time".

Table A13: Legibility measures by affective polarization, 2020 Election survey.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Within-party SD (σ)			Partisan gap (Δ)			Swamping (σ/D)			Party R^2		
	Low	Middle	High	Low	Middle	High	Low	Middle	High	Low	Middle	High
Unemployment rate (Sep 2020)	1.87	1.99	1.99	0.21	0.64	1.07	0.60	0.64	0.64	0.003	0.022	0.061
Border wall miles built	594	418	404	125	290	310	1.31	0.92	0.89	0.010	0.096	0.118
Biden tax plan ($60K family of four)	1,802	1,490	1,589	511	1,633	2,516	0.90	0.74	0.79	0.019	0.209	0.364

Notes. Affective polarization is a bipolar composite: for Democrats, (Biden positive emotions + Trump negative emotions) minus (Trump positive + Biden negative); sign-reversed for Republicans. Emotions are happiness, hope, and pride (positive) and anger, fear, and disgust (negative). Tertiles are equal-frequency splits of the composite within the pooled Democrat–Republican sample of the 2020 Election survey (N = 1,196 low, 1,389 middle, 905 high). All statistics are computed on errors (stated belief minus the truth benchmark). σ is the pooled within-party standard deviation of errors, Δ the absolute Democrat–Republican gap in mean errors, σ/D the swamping ratio, and party R^2 the share of error variance explained by party among Democrats and Republicans, as in Table 2. The swamping denominator D is each item's fixed policy magnitude and does not vary across tertiles.

Table A14: Native-survey reweighting robustness, with and without household income.

	(1)	(2)	(3)	(4)	(5)	(6)
	N	Unweighted σ/D	ANES-reweighted without income σ/D	ANES-reweighted with income σ/D	(4) − (2)	Effective N with income
Tax Cuts and Jobs Act (2017–18)						
Change in household tax burden	1,568	1.55	1.57	1.24	-0.316	264
2020 presidential election						
Biden tax plan ($60K family of four)	3,490	0.84	0.85	0.87	0.033	572
Border wall miles built	3,490	1.12	1.06	1.09	-0.033	572
Unemployment rate	3,490	0.63	0.62	0.62	-0.009	572
American Rescue Plan (2021)						
Household stimulus payment	2,617	0.68	0.70	0.77	0.087	687
Household net transfer	2,617	0.93	0.92	1.01	0.081	687
ARPA cost (% of federal spending)	2,862	1.02	1.04	1.04	0.023	754

Notes. The table recomputes the native-survey swamping ratios after raking each survey wave to ANES 2020 survey-weighted registered-voter margins. All estimates and target margins are restricted to Democrats and Republicans. Column (3) uses age group, sex, race/ethnicity, bachelor's-degree attainment, and party identification. Column (4) additionally uses household income in four bands: below $30,000; $30,000–$59,999; $60,000–$99,999; and $100,000 or more. Weights are computed separately by survey wave and capped at 10. For items with respondent-specific truth benchmarks, σ is the pooled within-party standard deviation of belief errors (belief minus truth); for common-truth items, σ is computed on stated beliefs. Weighted estimates center observations within party and pool squared deviations using the raking weights. Column (5) is column (4) minus column (2). N is the number of observations with nonmissing item, truth where applicable, and reweighting variables; effective N equals $(\sum_i w_i)^2 / \sum_i w_i^2$ for the income-inclusive weights.

Table A15: Robustness of noise measures to discretization of response scales.

	(1)	(2)	(3)	(4)	(5)
	Response format	Raw σ/Δ	Coarsened σ/Δ	Raw party R^2	Coarsened party R^2
Panel A. Tax Cuts and Jobs Act of 2017 (N = 1,568)					
Change in household tax burden	Free text	11.542	7.391	0.002	0.004
Share of cuts to top quintile	Slider	1.529	1.541	0.089	0.088
Share of cuts to bottom quintile	Slider	2.084	2.075	0.050	0.050
Panel B. 2020 Presidential Election (N = 3,490)					
Unemployment rate (Sep 2020)	Slider	4.720	4.658	0.011	0.011
Border wall miles	Slider	1.586	1.587	0.089	0.089
Biden tax plan ($60K family)	Slider	1.067	1.085	0.177	0.172
Panel C. American Rescue Plan Act of 2021 (N = 2,862)					
Household stimulus payment	Free text	7.011	7.552	0.005	0.004
Household net transfer	Free text	3.125	3.134	0.024	0.024
ARPA cost (% fed budget)	Free text	3.287	3.310	0.022	0.022

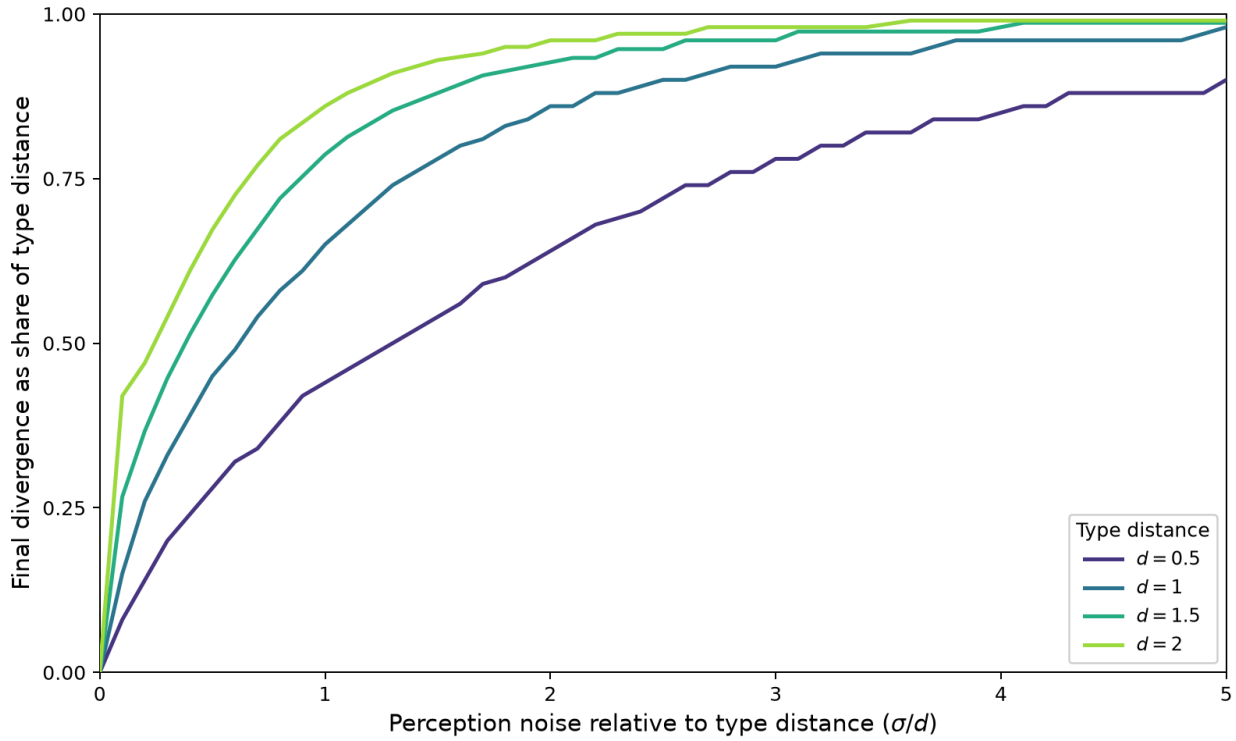
Notes. Responses are coarsened by assigning each stated belief to the nearest of 11 evenly-spaced anchor values spanning the observed response range. σ/Δ and party R^2 are computed on errors (stated belief minus external truth) among Democratic and Republican respondents, using the same conventions as Table 2.

Table A16: Calibrated candidate policy motivation ($\hat{\alpha}$) by issue and cycle.

	(1) 2016 ($\hat{\alpha}$)	(2) 2020 ($\hat{\alpha}$)	(3) 2024 ($\hat{\alpha}$)
Liberal–conservative	0.16	0.20	0.28
Government services vs. spending	0.13	0.14	0.16
Defense spending	0.09	0.12	0.10
Health insurance coverage	0.17	0.23	0.18
Jobs and standard of living	0.15	0.19	0.19
Aid to Black Americans	0.15	0.18	0.21
Environment vs. business	0.14	0.20	0.21
Abortion	0.18	0.21	0.25*
Median	0.15	0.19	0.20
Mean	0.15	0.18	0.20
Range	(0.09, 0.18)	(0.12, 0.23)	(0.10, 0.28)

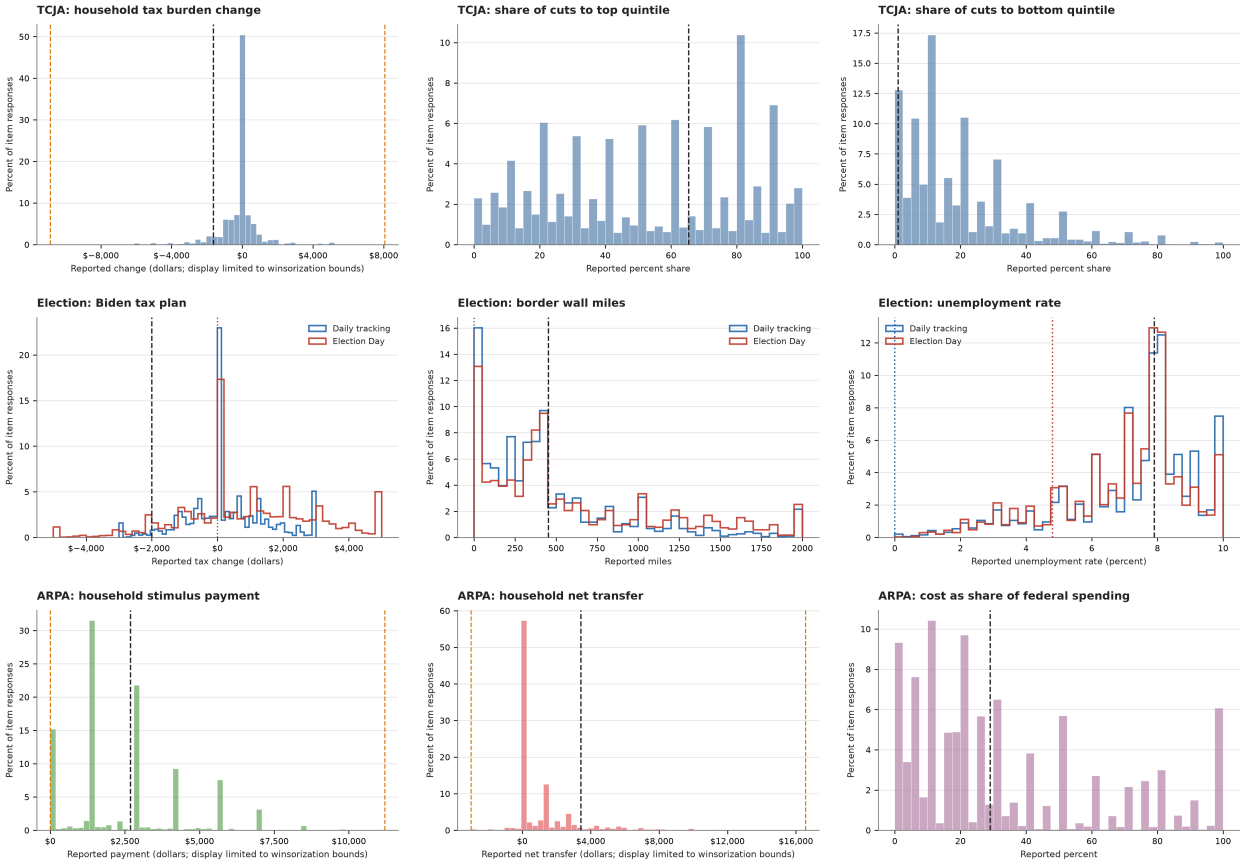
Notes. $\hat{\alpha}$ is the value of candidate policy motivation at which the simulation's median equilibrium divergence, evaluated at the item's observed swamping ratio σ/D, equals the target divergence implied by observed ANES candidate placements. The target is $D^{\text{target}} = (\Delta/(K-1)) \times 1.5$, where Δ is the mean absolute perceived gap between the Democratic and Republican candidates among Democratic and Republican respondents (survey-weighted) and K is the number of scale points. The median, mean, and range summarize across the eight issues within each cycle. *The 2024 abortion entry uses the 7-point placement item; 2016 and 2020 use the 4-point item.

Figure A1: Candidate divergence rises with perception noise at every candidate type distance



Notes. Final platform divergence as a share of the candidate type distance d (bliss points at $\pm d/2$), as a function of perception noise σ/d, holding candidate policy motivation α fixed at 0.175. Each line is a different d.

Figure A2: Item-level response distributions and elicitation defaults



Notes. Histograms show raw item responses, one row per survey. Black dashed lines mark each item's true value (for respondent-specific items, the sample mean of the respondent-specific statutory truth); orange dashed lines mark the cleaning-stage 0.5/99.5 percent winsorization bounds. Typed dollar-entry panels limit the display to the winsorization bounds while retaining all responses in percentage denominators. Election-survey panels distinguish the daily and Election Day instruments because the Biden-tax ranges and unemployment starting positions differed by wave; dotted lines mark documented slider starts. The TCJA household item and the ARPA items were typed numeric entries with no slider default; the TCJA quintile-share items were 0–100 sliders with no stored start position.

Appendix B. Native Survey Details

B1 Sample Recruitment and Composition

The native-survey evidence draws on original surveys fielded through Amazon Mechanical Turk via the CloudResearch platform between December 2017 and March 2021. The analytic sample contains 11,303 completed responses: 2,226 from the TCJA surveys, 4,861 from the 2020 election surveys, and 4,216 from the ARPA surveys. The 2020 election and ARPA samples retain respondents who reported being registered to vote; the TCJA samples, which did not consistently record registration, retain respondents who reported voting in the prior presidential election. Where stable worker identifiers are available, we keep the first completed, qualifying response from each worker. Such identifiers are unavailable in the 2020 election files. This appendix describes recruitment and summarizes the composition of the resulting samples.

The *TCJA* sample (N = 2,226) was assembled from two sub-waves fielded around the Tax Cuts and Jobs Act. The first coincided with the December 2017 Alabama special Senate election and was in the field from December 12 to December 19, 2017 (N = 860). The special election was on December 12, and the final version of the law was passed on December 15, making the events closely coincident. The second sub-wave coincided with the 2018 midterm election and was in the field from November 6 to November 13, 2018 (N = 1,366). These waves measured beliefs about the respondent's own change in household tax burden and the law's distributional incidence for the top and bottom of the income distribution. Respondents took a median of 8.2 minutes to complete the instrument (10.4 minutes in the December 2017 wave and 7.1 minutes in the November 2018 wave).

The *Election* sample (N = 4,861) was fielded around the 2020 general election. A daily tracking wave ran from October 16 to November 4, 2020 (N = 1,898), and an Election Day wave was collected on Election Day itself, November 3 (N = 2,963). These surveys measured beliefs about the September 2020 unemployment rate, the miles of new border

wall constructed, and what the Biden tax plan implied for a family of four earning $60,000. The median completion time was 8.0 minutes.

The *ARPA* sample (N = 4,216) was fielded between March 15 and March 27, 2021, immediately following passage of the American Rescue Plan Act. It comprised an initial wave, a second wave fielded several days later once the Act's stimulus payments had largely been disbursed (62 percent of its respondents had received a payment, versus 18 percent in the initial wave), and a supplementary wave recruiting only Republicans and Independents to enlarge the non-Democrat sample. Each survey measured beliefs about the respondent's own stimulus payment, own net transfer, and the cost of the Act as a share of 2020 federal spending. The median completion time was 6.3 minutes across the entire sample.

Table A1 reports the demographic composition of each sample alongside the pooled full sample. The three samples are broadly comparable. Respondents are in their late thirties to early forties on average, slightly under half male, roughly three-quarters white, and majority college-educated, with median household income in the high forties to low fifties of thousands of dollars. Two demographic items—marital status and presence of children—were not collected in the Election wave and are reported only where available.

The final column benchmarks the pooled sample against the composition of registered voters in the November 2020 Current Population Survey (Fabina and Scherer, 2022). As is typical of online convenience samples, our respondents differ from the registered-voter population on several observable dimensions. Sixty-four percent of the pooled sample holds a bachelor's degree, against 40 percent of registered voters, and respondents are roughly a decade younger on average (39.6 versus approximately 50 years). They are also modestly whiter (73 versus 70 percent) and slightly less male (46 versus 47 percent). Despite their higher education, respondents report lower household income than the country as a whole—a median of $47,500 against the $67,521 national median for all U.S. households in 2020 (Shrider et al., 2021). Because composition could affect within-party dispersion as well as population means, Table A14 reports two reweighting specifications. Without income, all

seven ratios change by at most 4.8 percent. With income, the TCJA household-tax ratio falls from 1.55 to 1.24, the ARPA stimulus-payment ratio rises from 0.68 to 0.77, and the ARPA net-transfer ratio rises from 0.93 to 1.01; the rank ordering of the seven items is unchanged. The income-inclusive effective sample sizes are 264, 572, and 754 in TCJA, Election, and ARPA. The exercise does not convert the native surveys into probability samples and cannot address unobservable selection into online survey participation.

Table A2 turns to politics, benchmarked against the weighted ANES 2020 Time Series (American National Election Studies, 2021). All retained Election and ARPA respondents report being registered, while all retained TCJA respondents report voting in the prior presidential election; 81 percent of the Election sample recalls voting in 2016. Registration was not consistently recorded in the TCJA instruments, and prior-election turnout was not collected in ARPA, so we do not pool these measures across all three episodes. Relative to the ANES electorate, the sample tilts modestly left. Democrats are overrepresented (41 versus 35 percent), and party feeling thermometers run warmer toward the Democratic Party (52 versus 45) and cooler toward the Republican Party (36 versus 45). These compositional differences motivate the reweighting exercise above.

Because the TCJA and ARPA instruments asked respondents to reason about their own tax situation, both collected the household tax information needed to construct each respondent's true welfare consequence. Table A3 summarizes household adjusted gross income, the number of dependents, household size, and tax-filing status. In both samples, married-filing-jointly and single filers each account for roughly 45 percent of respondents, with head-of-household and married-filing-separately filers making up the remainder.

B2 Belief Elicitations

B2.1 Tax Cuts and Jobs Act of 2017

1. **Direction of change in personal federal tax burden** Participants are asked to report how their 2018 federal tax burden will compare to their 2017 federal tax burden

as higher, lower, or unchanged. The responses were coded as 1, -1, and 0, respectively.

2. **Change in personal federal tax burden**. If participants reported believing that their federal tax burden would be higher or lower in 2018 than 2017, we asked participants to report by how much their taxes would change in dollars, rounding to the nearest hundred dollars.

3. **Share of tax cuts to top and bottom income quintiles**. Participants use a slider ranging from 0 to 100 to report their belief about the share of tax cuts that go to the poorest and richest 20 percent of Americans.

B2.2 2020 Presidential Election

1. **Change in unemployment rate during the Trump administration**. Participants are told that when Donald Trump took office in January 2017, the unemployment rate was 4.8 percent. They are asked to report their belief about the unemployment rate as of September 2020 using a slider that ranges from 0 to 10. We difference the report from 4.8 to create a measure of participants' beliefs about the change in unemployment over this period.

2. **Miles of new border wall constructed during the Trump administration**. Participants are told that the border between Mexico and the United States is approximately 2,000 miles long. They are asked to report their belief about how many miles of new border wall were constructed during the Trump administration using a slider from 0 to 2,000 miles.

3. **Proposed Biden tax change for a typical family**. Participants were asked to report their belief about how much more or less a typical family earning $60,000 per year would pay in taxes under Joe Biden's proposed tax plan. The daily wave used a slider ranging from −$3,000 to $3,000; the Election Day wave used −$5,000 to $5,000.

B2.3 American Rescue Plan Act of 2021

1. **Eligibility for direct stimulus payment**. Participants were told that ARPA will disburse direct payments to some Americans. They were then asked to enter the total dollar amount that their household would receive from the law. If they did not expect to receive a direct payment, they were told to enter zero. Participants who entered a nonzero value are coded as believing they are eligible for a payment (1), while those who entered zero are coded as believing they are not eligible (0).

2. **Direction of change in federal tax burden from ARPA**. Participants were told, "The American Rescue Plan may affect your finances this year both through a direct stimulus payment and changes to your federal taxes for 2021." They were then asked to report their belief as to whether ARPA would increase, decrease, or leave their personal federal tax burden unchanged (separately from the direct payment), leaving all else equal. The responses were coded as 1, -1, and 0, respectively.

3. **Total size of net transfer from ARPA**. Participants were told to enter the total net transfer they would receive as a result of ARPA. Positive values signify payments from the government to participants' households, where negative values reflect net payments from the participants' household to the government.

4. **Cost as percentage of 2020 federal budget**. Participants were asked to report their belief about the cost of ARPA as a percentage of the 2020 federal budget as a value between 0 and 100.

B3 Processing and Elicitation Details

All data were collected on Amazon Mturk via the CloudResearch platform. Where a stable MTurk worker identifier is available, we retain only the first completed, qualifying response from each worker. Stable worker identifiers are unavailable in the 2020 election files, so deduplication cannot be applied there.

Three open-ended household-dollar belief variables are winsorized during cleaning, including the TCJA change in household tax burden (`belief_tax`), the ARPA direct payment (`belief_direct`), and the ARPA net transfer (`belief_net`). For each variable, values below the 0.5th percentile are set to that percentile and values above the 99.5th percentile are set to that percentile. The resulting variables (`belief_taxw`, `belief_directw`, and `belief_netw`) enter every reported analysis of those items, including the decomposition and noise-measure tables, headline swamping comparison, reweighting and heterogeneity exercises, bootstrap standard errors, discretization and completion-time checks, and the reliability-bound analysis. Bounded election-survey items and the ARPA cost item are not winsorized, and neither are any statutory or common-truth benchmarks. Figure 1 applies a separate 5th–95th percentile display trim within party, as stated in its note.

The ARPA net-transfer sample is restricted before winsorization to raw reports between −100,000 and 100,000 dollars. This is a distinct plausibility screen rather than part of the percentile rule. The baseline ARPA direct-payment item has no analogous dollar screen. Appendix Table A5 holds each denominator and item sample fixed and replaces the three winsorized beliefs with their preserved raw values; it also reports a common $|x| \leq 100{,}000$ dollar diagnostic. Appendix Figure A2 displays item-level response distributions. Removing winsorization raises the TCJA ratio from 1.55 to 2.21 and the ARPA net-transfer ratio from 0.93 to 1.24. The fully raw ARPA direct-payment ratio is dominated by a $219,000,500 entry; under the common plausibility screen it rises from 0.68 to 1.63. Bounded-item estimates are unchanged.

The three 2020 election factual items were sliders. The border-wall slider ranged from 0 to 2,000 and started at 0 in both waves. The unemployment slider ranged from 0 to 10; it started at 0 in the daily wave and at 4.8, the baseline stated in the question, in the Election Day wave. The Biden-tax slider started at 0 in both waves, with the wave-specific ranges stated above. The border-wall and unemployment sliders required a response in both waves. The Biden-tax response was optional in the daily wave and required in the Election

Day wave. By contrast, the headline TCJA and ARPA household-dollar beliefs and the ARPA cost item were typed numeric entries, so they had no slider starting position; the ARPA entries were required. The archived Qualtrics instruments and a machine-readable extraction of these settings are included in the replication materials.

Appendix C. Estimates of True Values For Each Policy Provision

C1 Tax Cuts and Jobs Act of 2017

1. **Change in personal federal tax burden**. We estimate each respondent's federal tax burden under projected 2018 prior law and under the TCJA. The calculator applies filing-status-specific brackets, standard deductions, prior-law personal exemptions and their phaseout, the Child Tax Credit and its statutory phaseouts and refund limits, and the Earned Income Tax Credit. Nonrefundable CTC amounts are applied before the Additional Child Tax Credit refund ceiling, and the independently refundable EITC is then subtracted. Income, filing status, and dependent counts come from the survey; income brackets are represented by their midpoints. We calculate the change as the TCJA burden minus the prior-law burden using the following steps.

   (a) Subtract the standard deduction from the participants' pretax income under both regimes.

   (b) Calculate the participant's income tax burden under both regimes.

   (c) Calculate the value of the Child Tax Credit under both regimes, accounting for its partial refundability.

   (d) Calculate the value of the Earned Income Tax Credit under both regimes.

   (e) Compute a final federal tax burden under both regimes by subtracting the value of the tax credits from the income tax burden.

   (f) Compute the change in tax burden by subtracting the prior burden from the burden under the TCJA rules.

   The code that reproduces these calculations is available in the replication materials.

2. **Direction of change in personal federal tax burden**. Given the above calculations, participants are assigned a value of -1 if their taxes are lower under TCJA, 0 if their taxes are unchanged under TCJA, and 1 if their taxes go up under TCJA.

3. **Share of tax cuts to top and bottom income quintiles**. Estimates of the true share of cuts are drawn from a Tax Policy Center report entitled "Effects of the Tax Cuts and Jobs Act, A Preliminary Analysis."

C2 2020 Presidential Election

1. **Change in unemployment rate during the Trump administration**. The unemployment rate estimate of 7.9 percent was drawn from a Bureau of Labor Statistics report issued on October 2, 2020 (Bureau of Labor Statistics, 2020).

2. **Miles of new border wall constructed during the Trump administration**. We use 453 miles, the construction total reported in the U.S. Customs and Border Protection *Border Wall Status* memorandum dated January 8, 2021 (U.S. Customs and Border Protection, 2021). The relevant estimate appears in the table on the memorandum's first page.

3. **Proposed Biden tax cut for typical family**. We construct a Child Tax Credit-only benchmark for a married family earning $60,000 with two qualifying children ages 6–16. Under then-current law, the maximum credit was $2,000 per qualifying child; Biden proposed increasing it to $3,000 per qualifying child in this age range (Mermin, Rohaly and Nunns, 2020). Holding other provisions fixed, the two $1,000 increases imply a $2,000 reduction in federal tax liability. This is an explicitly constructed family benchmark, not the $680 income-group average reported in Table 2 of the Tax Policy Center analysis.

C3 American Rescue Plan Act of 2021

1. **Eligibility for direct stimulus payment**. We calculate the third-round economic-impact payment from each respondent's reported adjusted gross income, filing status, and household size. The full payment is $1,400 per household member. It phases out

linearly over the statutory AGI windows: $75,000–$80,000 for single and married-filing-separately filers, $112,500–$120,000 for heads of household, and $150,000–$160,000 for joint filers. Positive calculated payments are coded as eligible. Respondents missing filing status or other required household inputs have no respondent-specific benchmark but remain eligible for the common-truth cost analysis.

2. **Change in personal federal tax burden**. We calculate 2021 income tax and the Child Tax Credit under pre-ARPA and ARPA rules, including filing-status-specific phaseouts and the pre-ARPA refundable-credit limit, and compare the 2021 EITC schedule with the schedule that would have applied absent ARPA. Because earned income is not elicited separately, adjusted gross income is used as its documented proxy for refundable-credit phase-ins. The survey records the number of qualifying children but not their ages, so the main benchmark conservatively treats all qualifying children as ages 6–16 (a $3,000 ARPA credit); the calculator also retains an all-under-six $3,600-per-child upper bound. The reported main results use the conservative benchmark.

3. **Total size of net transfer from ARPA**. From the above two calculations, we sum the direct payments and reduction in tax burden to arrive at the total size of the ARPA net transfer.

4. **Cost as percentage of 2020 federal budget**. The total cost of ARPA was $1.9 trillion, according to a US Department of the Treasury fact sheet (U.S. Department of the Treasury, 2022). The 2020 federal budget was $6.6 trillion, according to the Congressional Budget Office (Congressional Budget Office, 2020). Dividing the two values (and rounding, to account for the slider units available to participants) gives our estimate of 29%.

C4 Reliability Bound for Reported Household Inputs

Let τ_i^* denote the true household policy consequence and let the recorded statutory benchmark be $\tau_i = \tau_i^* + u_i$, where u_i is classical measurement error. If respondents know their true consequence perfectly, an OLS regression of stated beliefs on the recorded benchmark has slope

$$\lambda = \frac{\mathrm{Var}(\tau_i^*)}{\mathrm{Var}(\tau_i^*) + \mathrm{Var}(u_i)},$$

the reliability of the benchmark. More generally, if stated beliefs load on the true consequence with slope γ, the observed belief–truth slope is $\gamma\lambda$. The variances in this expression are understood after residualizing the benchmark on party, matching the party fixed effects in the reported regressions.

Table C1 uses this relationship to bound the role of benchmark error. The ARPA stimulus payment is a conservative validation item because its statutory value is a relatively simple function of the same reported household inputs. Attributing its entire attenuation from one to benchmark error gives reliability of 0.609, or a benchmark-error variance share of 39.1 percent. By contrast, explaining the TCJA slope of 0.083 with accurate beliefs requires reliability of only 0.083, so about 92 percent of recorded-benchmark variance would have to be measurement error. If reliability is at least the stimulus-item benchmark, the corrected belief–truth slope is at most 0.14 for TCJA and 0.41 for the ARPA net transfer. Thus plausible input error can attenuate the estimated relationship, especially for the net-transfer item, but cannot rationalize the near-zero TCJA slope. This exercise bounds rather than eliminates benchmark error because true household inputs are not independently observed. Common-truth items do not use respondent-specific household inputs and are unaffected.

Table C1: Reliability bounds for respondent-specific truth benchmarks.

	(1) Sample size (N)	(2) Belief–truth slope	(3) 95% confidence interval	(4) Required reliability	(5) Required benchmark-error variance share	(6) Maximum corrected slope
Change in household tax burden (TCJA)	2,170	0.083	[0.055, 0.110]	0.083	91.7%	0.14
Household stimulus payment (ARPA)	3,735	0.609	[0.583, 0.636]	0.609	39.1%	1.00
Household net transfer (ARPA)	3,735	0.249	[0.224, 0.275]	0.249	75.1%	0.41

Notes. Column (2) reports the slope from a regression of stated belief on the respondent-specific statutory truth benchmark with party fixed effects; column (3) reports the associated 95% confidence interval. N is the complete-case sample with nonmissing belief, benchmark, and party. Under classical measurement error, if respondents know their true policy consequence perfectly, the observed slope equals the reliability of the recorded benchmark; columns (4) and (5) report this required reliability and the corresponding benchmark-error variance share. Column (6) divides each observed slope by the slope for the ARPA stimulus payment and caps the ratio at one.

Appendix D. Robustness Checks for Characterizing Noise in Beliefs

D1 Uniform-Response Benchmarks

Table 3 asks how much dispersion would arise from independent uniform responses on each item's elicitation scale. Unlike a simulation calibrated to the observed distribution, this exercise does not preserve the observed mean or variance. Instead, it supplies a scale-specific reference point. Each observed dispersion-to-signal ratio is compared with the ratio generated by uninformative use of that item's own response scale.

Observed dispersion in candidate placements is well below uniform-response benchmarks. The native placement ratio is 0.46 of its benchmark, and the ANES placement rows are 0.53 to 0.55. The four bounded policy and factual items for which the benchmark can be constructed lie closer to uniform-response dispersion, ranging from 0.67 for the Biden-tax item to 1.02 for ARPA cost. Thus, among bounded items, the placement–policy asymmetry is not mechanically generated by response-scale width or granularity.

D2 Calculating Legibility Measures for Candidate Placements

D2.1 Swamping Ratio for Candidate Placements

Let $y_{i,D}$ and $y_{i,R}$ denote respondent i's placement of the Democratic and Republican candidates, and let

$$S_i = y_{i,R} - y_{i,D}$$

denote the respondent's perceived candidate separation. The distance between candidates is

$$D_{placement} = \frac{1}{N} \sum_i |S_i|,$$

the mean absolute perceived separation, constructed exactly as D_q^{policy} is constructed for respondent-specific welfare items in Section 4.3, with S_i in place of the true effect τ_i.

The numerator pools within-party dispersion across both candidates as well as both

parties.

$$\sigma = \sqrt{\frac{(n_D - 1)(s^2_{D,D} + s^2_{D,R}) + (n_R - 1)(s^2_{R,D} + s^2_{R,R})}{2(n_D - 1) + 2(n_R - 1)}},$$

where $s^2_{g,c}$ is the sample variance of $y_{i,c}$ among respondents with party g, each candidate centered on its own party-specific mean before pooling. Centering separately by candidate keeps the between-candidate gap out of σ; pooling within party keeps projection bias (Democrats and Republicans perceiving a given candidate differently), out of it as well.

$\sigma / D_{placement}$ can then be interpreted as in Section 4.3. A swamping ratio of one indicates that within-party dispersion in candidate placements is the same magnitude as the separation between them.

D2.2 Denominator Comparability and Candidate Ordering

The baseline denominator $D_{placement}$ is perceptual because the electoral decision problem depends on the candidate difference voters believe they face. An exclusively objective denominator would answer a different question and would combine two sources of a small perceived gap, namely actual platform convergence and imprecision in candidate placement. Unlike D_q^{policy}, however, $D_{placement}$ is not an external truth benchmark. We therefore report two diagnostics.

First, we calculate the share of paired placements that reverses the expected candidate ordering on each scale, with ties reported separately. The expected direction follows the response coding and the candidates' party positions. The Republican candidate is expected at the lower-coded endpoint for government services versus spending and the four-category abortion item, and at the higher-coded endpoint for the remaining scales.

Second, let $W_q = K_q - 1$ denote the width of item q's placement scale. For an externally specified candidate separation equal to share f of the scale, we replace the perceived denominator with

$$D_q^{\text{external}}(f) = fW_q$$

and calculate

$$R_q^{\text{external}}(f) = \frac{\sigma_q}{fW_q}.$$

This keeps each item's observed within-party placement dispersion fixed while varying only the denominator. We report $f \in \{0.20, 0.25, 1/3, 0.50\}$ and calculate the common break-even value at which the median placement ratio equals the policy-item median.

Table A10 reports the results. Median perceived separation is 55.0 to 62.3 percent of scale width across the three placement groups. At an external separation of one-third of scale width, the median placement ratio is 0.788, compared with the policy median of 0.930. The break-even separation is 28.2 percent of scale width. External separations above that value preserve the placement–policy ordering; smaller values attenuate or reverse it.

D2.3 Party R^2 for Candidate Placements

Party R^2 for candidate placements pools sums of squares across both candidates rather than averaging two separate R^2 values. Each respondent supplies two placements, and averaging ratios with different denominators can misstate the combined explanatory share. For candidate $c \in \{D, R\}$, let

$$SS_c^{\text{tot}} = \sum_i (y_{i,c} - \bar{y}_c)^2, \qquad SS_c^{\text{win}} = \sum_{g \in \{D,R\}} \sum_{i:g_i=g} (y_{i,c} - \bar{y}_{g,c})^2$$

denote the total and within-party sums of squares for candidate c, where $\bar{y}_{g,c}$ is the mean placement of candidate c among respondents with party g and $\bar{y}_c$ is the overall mean. Party R^2 for the placement item is

$$R^2 = 1 - \frac{SS_D^{\text{win}} + SS_R^{\text{win}}}{SS_D^{\text{tot}} + SS_R^{\text{tot}}}.$$

Independents are excluded from this calculation.

D3 Measurement of Heterogeneity Variables

D3.1 Political Sophistication (ANES 2016, 2020, 2024)

Political sophistication is constructed from a four-item civic knowledge battery administered in each ANES pre-election survey. Respondents are scored 0 (incorrect or missing) or 1 (correct) on each item, and the four scores are summed to a 0–4 knowledge index. Low sophistication is defined as zero or one correct answer and high sophistication as three or four correct answers; respondents with exactly two correct answers are excluded. ANES sampling weights (V160101 in 2016, V200010b in 2020, and V240003 in 2024) are used throughout.

The four knowledge items are identical in substance across cycles, though ANES variable identifiers differ.

Item 1. Senator term length. "For how many years is a United States Senator elected—that is, how many years are there in one full term of office for a U.S. Senator?" Participants provide a number. The correct answer is six. ANES variable IDs are V161513 (2016), V201644 (2020), V241612 (2024).

Item 2. Least-funded federal program. "On which of the following does the U.S. federal government currently spend the least?" Response options include Foreign aid; Medicare; National defense; Social Security. The correct answer is foreign aid. ANES variable IDs are V161514 (2016), V201645 (2020), V241613 (2024).

Item 3. Pre-election House majority. "Do you happen to know which party currently has the most members in the U.S. House of Representatives in Washington?" Response options include Democrats; Republicans. The correct answers are Republicans (2016 and 2024); Democrats (2020). ANES variable IDs are V161515 (2016), V201646 (2020), V241614 (2024).

Item 4. Pre-election Senate majority. "Do you happen to know which party currently has the most members in the U.S. Senate?" Response options include Democrats; Republicans. The correct answers are Republicans (2016 and 2020); Democrats (2024). ANES variable IDs are V161516 (2016), V201647 (2020), V241615 (2024).

Analysis split. The knowledge score is the number of correct answers on the four ANES civics items (length of a full Senate term, the program with the least federal spending, and pre-election House and Senate majorities), computed for respondents answering all four. Low sophistication is zero or one correct; high sophistication is three or four correct; respondents with exactly two correct are omitted. Thresholds are fixed across survey cycles. All statistics use ANES sampling weights.

D3.2 Self-Reported Attention to National Politics (Native Surveys)

Each of the three native survey waves included a self-reported attention item. The question wording and response scale differ across waves, so splits are defined wave-specifically on a common high/lower distinction.

TCJA wave (December 2017 + November 2018). "How often did you pay attention to politics during this special election?" The response options and retained-sample counts are all the time (1; $n = 507$), most of the time (2; $n = 992$), some of the time (3; $n = 274$), rarely or never (4; $n = 61$), and about half the time (5; $n = 392$). Because the raw codes are not monotonic, we define high-attention respondents by their labels—all or most of the time (codes 1 or 2)—and all others as lower attention.

2020 Election wave (October–November 2020). "How much attention would you say you've paid to the 2020 presidential election? Please indicate your answer on a scale from 1 (no attention) to 10 (an extreme amount of attention)." Pooled across the daily and Election Day waves, retained-sample counts are 1 ($n = 31$), 2 ($n = 66$), 3 ($n = 137$), 4 ($n = 170$), 5 ($n = 458$), 6 ($n = 414$), 7 ($n = 774$), 8 ($n = 1{,}035$), 9 ($n = 942$), and 10 ($n = 828$), with six

missing responses. We define high attention as 8–10 and all other nonmissing responses as lower attention.

ARPA wave (March 2021). "How often would you say you pay attention to national politics?" The five response options follow.

1. I paid attention rarely or never
2. I paid attention some of the time
3. I paid attention about half the time
4. I paid attention most of the time
5. I paid attention all the time

Value counts in the retained ARPA sample are rarely or never ($n = 88$), some of the time ($n = 582$), about half the time ($n = 821$), most of the time ($n = 1{,}889$), and all the time ($n = 836$). We recode these labels monotonically from 1 to 5 and define high-attention respondents as those answering "most of the time" or "all the time" (recoded 4 or 5). All others are defined as lower attention.

Affective Polarization (2020 Election Survey)

Affective polarization is measured using a set of 12 candidate emotion items administered in the 2020 Election survey. Respondents rated how strongly each of six emotions described their feelings toward each candidate (Donald Trump and Joe Biden). Each emotion is recorded on a 3-point scale with 1 = not at all; 2 = somewhat; 3 = very much.

The six emotions are anger, happiness, fear, hope, disgust, and pride. From these 12 items a bipolar affective-polarization composite is constructed separately for Democrats and Republicans. For *Democrats*:

$$\begin{aligned}\text{Composite}_{\text{Dem}} = &\underbrace{(\text{Biden happiness} + \text{Biden hope} + \text{Biden pride})}_{\text{in-party positive}} \\ &+ \underbrace{(\text{Trump anger} + \text{Trump fear} + \text{Trump disgust})}_{\text{out-party negative}} \\ &- \underbrace{(\text{Trump happiness} + \text{Trump hope} + \text{Trump pride})}_{\text{out-party positive}} \\ &- \underbrace{(\text{Biden anger} + \text{Biden fear} + \text{Biden disgust})}_{\text{in-party negative}}\end{aligned}$$

For Republicans the sign is reversed (Trump is the in-party candidate, Biden the out-party candidate). Respondents identifying as independent are excluded from this analysis. All six positive and all six negative emotions must be nonmissing for a respondent to receive a composite score; respondents with any missing emotion item are excluded. The composite is divided into equal-frequency tertiles within the pooled Democrat–Republican sample. The top tertile is labeled "high affect"; the bottom, "low affect."

Appendix E. Model and Simulation Implementation

E1 Model and Simulation Details

We solve the model by iterating candidate best responses on a discrete platform grid spanning $[-1.5, 1.5]$ at increments of 0.01. In each iteration, each candidate takes the opponent's current platform as given and selects the grid point that maximizes her payoff. We repeat for up to 50 cycles or until neither candidate's platform moves by more than half a grid step, whichever comes first. To characterize how outcomes vary with the draw of voter ideal points, we run 100 independent replications, each drawing $N = 1{,}000$ ideal points from $\theta_i \sim U[-1, 1]$, and report median final platform distance. For every grid point with $\sigma/d^{type} \geq 0.2$, all 100 baseline replications converge under each reported value of α; at $\sigma/d^{type} = 0.1$, at least 94 of the 100 replications converge. At exactly zero noise, the discontinuous deterministic best responses cycle and none meets the convergence criterion; the stored zero-noise value is the terminal distance after 50 cycles. We therefore use positive-noise points for quantitative comparisons.

E2 Observed and Uniform-Response Dispersion on the Model Scale

The observed-to-uniform ratio from Table 3 is a scale-normalized descriptive statistic, not a direct value of the model parameter σ/d. Setting σ/d equal to that share would force independent uniform responding to equal one, even though the Gaussian model reaches uninformative perception only in the limit as $\sigma/d \to \infty$. We therefore retain the model's existing units and separately map the observed σ/D and the scale-specific uniform-response σ/D to the calibrated curve. Table E1 reports the resulting comparison at $\alpha = 0.175$. The exercise is limited to bounded items with a natural uniform-response benchmark.

Table E1: Model divergence at observed and uniform-response dispersion.

	(1) Observed σ/d	(2) Uniform-response σ/d	(3) Observed / uniform	(4) Observed model divergence	(5) Uniform-response model divergence	(6) Difference (p.p.)
Candidate placements						
Native placement [1–7]	0.40	0.88	0.46	51.5%	74.5%	23.0
ANES placements [1–7]	0.48	0.88	0.55	56.1%	74.5%	18.4
ANES placements [1–4]	0.48	0.89	0.53	55.9%	75.1%	19.3
Bounded policy and factual beliefs						
ARPA cost [0–100]	1.02	1.00	1.02	79.1%	78.5%	-0.6
Biden tax plan [±$3,000/$5,000]	0.84	1.25	0.67	73.4%	84.3%	10.9
Border wall miles [0–2,000]	1.12	1.27	0.88	81.7%	84.8%	3.1
Unemployment rate [0–10]	0.63	0.93	0.68	64.2%	76.4%	12.2

Notes. Columns (1) and (2) place the observed dispersion and the scale-specific independent-uniform-response benchmark on the model's existing σ/d axis. Column (3) is their ratio. Columns (4) and (5) report equilibrium platform divergence as a share of the candidate bliss-point distance at $\alpha = 0.175$; column (6) is the percentage-point difference. The uniform-response value is a survey-scale reference point, not the $\sigma/d \to \infty$ pure-noise limit of the Gaussian model. Open-ended household-dollar items have no uniform-response benchmark and are omitted.

E3 Directional and Party-Block Perception Bias

The baseline model isolates independent, mean-zero perceptual dispersion. To test whether its comparative static depends on that restriction, we augment perceived platforms with a systematic component,

$$\tilde{x}_{ij} = x_j + b_{g(i)} + \eta_{ij}, \qquad \eta_{ij} \sim N(0, \sigma^2),$$

where $g(i)$ denotes voter i's party block. The same $b_{g(i)}$ shifts both candidates for a voter, so it leaves perceived candidate ordering unchanged but shifts the perceived midpoint. Residual candidate-specific errors remain independent. This specification captures the empirically important correlation between party and directional perception error without changing the mechanism through which residual dispersion compresses vote-share gradients.

We calibrate $b_{g(i)}$ using errors on the Biden tax-plan item, the closest survey analogue to perceptions of competing candidate platforms. The overall mean error is $0.687d^{type}$. Party-specific means are $0.993d^{type}$ for Democrats, $0.781d^{type}$ for Independents, and $0.204d^{type}$ for Republicans. Subtracting the overall mean gives centered party-block shifts of $0.306d^{type}$, $0.094d^{type}$, and $-0.483d^{type}$. We assign model voters to ideologically ordered Democratic,

Independent, and Republican blocks using their empirical shares in this item (42.1, 26.0, and 31.8 percent). The common-bias specification sets every $b_{g(i)}$ to the overall mean; the party-block specification uses the centered shifts; and the combined specification uses the party-specific means.

Table E2 reports the results for $\alpha = 0.175$. From $\sigma/d^{type} = 0.5$ through 2.0, equilibrium divergence is weakly increasing at every adjacent grid point under the baseline, common-bias, party-block, and combined specifications; the rank correlation is 1.00 in each case. At $\sigma/d^{type} = 0.9$, the corresponding divergence shares are 0.753, 0.800, 0.813, and 0.840. Directional and party-linked bias shift equilibrium midpoints and raise divergence levels, but the central comparative static survives. Increasing residual dispersion continues to weaken moderation incentives. Across this range, convergence is at least 96 percent for every specification. Below 0.5, the common-bias and combined-bias models frequently cycle under sequential best responses, so we do not use those low-noise endpoints to evaluate monotonicity.

Table E2: Model robustness to directional and party-block perception bias.

Panel A. Error-structure calibration

Error structure	Common shift (b/d)	Dem. block deviation	Ind. block deviation	Rep. block deviation	Left vote share at fixed types, $\sigma/d = .9$
Mean-zero independent	+0.000	+0.000	+0.000	+0.000	0.500
Common directional bias	+0.687	+0.000	+0.000	+0.000	0.682
Party-block bias	+0.000	+0.306	+0.094	-0.483	0.510
Common + party-block bias	+0.687	+0.306	+0.094	-0.483	0.642

Panel B. Equilibrium divergence as a share of candidate-type distance

Error structure	$\sigma/d = .5$	$\sigma/d = .9$	$\sigma/d = 1.5$	$\sigma/d = 2$	Spearman correlation	Minimum convergence
Mean-zero independent	0.573	0.753	0.880	0.927	1.000	1.00
Common directional bias	0.640	0.800	0.893	0.933	1.000	0.99
Party-block bias	0.713	0.813	0.893	0.933	1.000	1.00
Common + party-block bias	0.780	0.840	0.907	0.940	1.000	0.96

Notes. Bias moments are calibrated from the Biden tax-plan item, the closest empirical analogue to competing platforms. Its overall mean error is $0.687D$; party-specific mean errors are $0.993D$ for Democrats, $0.781D$ for Independents, and $0.204D$ for Republicans. The party-block rows subtract the overall mean so their weighted average is zero. Model voters are assigned to ideologically ordered party blocks with empirical shares 42.1%, 26.0%, and 31.8%. The bias shifts both perceived platforms within a voter block and therefore shifts the perceived midpoint; residual candidate-specific errors remain independent normal draws with standard deviation σ. Panel A's vote share holds platforms at the candidate bliss points. Panel B reports medians among converged runs across 100 voter draws at $\alpha = .175$. Spearman correlations and minimum convergence rates use the $\sigma/d = .5$ to 2 grid in increments of .1. Divergence is weakly increasing at every adjacent point for all four structures. Below .5, the common-bias and combined-bias specifications frequently cycle under sequential best responses, so those points are not used in the comparative-static test.