Accountability Among Allies:

Intra-Party Competition and the Electoral Cost of Scandal

Stefano Sangiovanni[*§] Giovanna Maria Invernizzi[†§] Andrea Ceron[‡]

September 9, 2026

Abstract

When do voters electorally punish politicians for scandals? We argue that a scandal changes votes only when it can tip voters who are near indifference between candidates, and that intra-party contests place far more voters in that position than inter-party contests. We formalize this argument in a model of electoral sanctioning and test it in a pre-registered conjoint experiment on 3,000 U.S. adults. The scandal penalty is about 57% larger in intra-party contests, and, as the model predicts, the amplification appears even when both candidates come from the party the voter opposes. Inter-party choice settings may therefore understate how strongly voters punish misconduct when credible co-partisan alternatives are available.

Keywords: intra-party competition; electoral sanctioning; political scandals; candidate substitutability; factions; primaries; conjoint experiments.

[*]University of Milan. E-mail: stefano.sangiovanni@unimi.it
[†]Bocconi University. E-mail: giovanna.invernizzi@unibocconi.it. Corresponding author.
[‡]University of Milan. E-mail: andrea.ceron@unimi.it

1 Introduction

> "Follow me around. I don't care. I'm serious. If anybody wants to put a tail on me, go ahead. They'll be very bored."
>
> Gary Hart, during the 1988 Democratic presidential nomination campaign (Dionne 1987)

In May 1987, Gary Hart was the clear front-runner for the Democratic presidential nomination. Within a week of press reports about an extramarital affair, his campaign had collapsed (Dionne 1987). We cannot know what the same revelations would have done to him in a general election. But the episode points to a possibility that has gone largely untested, and that is the subject of this paper: punishing misconduct may be easier when the alternative is a co-partisan.

A large body of research asks whether voters sanction political misconduct. Reviewing it, De Vries and Solaz (2017) conclude that voters do punish corrupt politicians, but far less consistently and less severely than models of electoral accountability would lead one to expect, and that how much they punish depends heavily on context. The standard explanation is partisanship. Voters discount damaging information about their own side and weigh it heavily against the other, a pattern consistent with more general accounts of motivated reasoning about politics (Taber and Lodge 2006; Lee et al. 2023), and evidence that partisan information can crowd out other candidate considerations (Kirkland and Coppock 2018). Graham and Svolik (2020) give this logic its sharpest formulation: when voters must trade valence considerations against partisan interests, most act as partisans first. In their candidate-choice experiment, the typical American is unwilling to cross party lines even to punish violations of democratic principles.

This diagnosis carries an implication that has not been tested directly. If weak sanctioning reflects the price of crossing party lines, then sanctioning should become easier when that price is removed—when the alternative to a scandalized candidate shares the candi-

date's party. This choice structure appears wherever a party's own supporters choose among its candidates: primaries and other nomination contests, leadership elections (Loewen and Rubenson 2011), and factional challenges in which an organized bloc runs a rival against the incumbent leadership. What matters for our argument is that every option on the ballot carries the same party label. These are not marginal settings: in the United States especially, many districts are safe for one party, so the decisive contest may occur before the general election, where intra-party accountability shapes who actually holds office.

Yet the comparison this implication calls for is largely missing from existing evidence. Nearly all experimental estimates of scandal effects come from settings where party is an attribute of a single candidate, not of the choice set (Incerti 2020). Whether the same scandal carries a different penalty when the alternative is a co-partisan rather than an out-partisan therefore remains, to our knowledge, unanswered. Answering it requires randomizing the composition of the choice set—a step reviews of the corruption-voting literature identify as necessary but difficult, since the alternatives a voter faces are ordinarily endogenous to the misconduct she is asked to punish (De Vries and Solaz 2017).

We provide that evidence, and a simple theoretical account of why the answer matters. We formalize the mechanism in a model of electoral sanctioning in the random-utility tradition (McFadden 1974; Train 2009), treating scandal as a negative valence attribute traded off against party, faction, and other candidate characteristics. The model delivers two main implications. First, a scandal costs the scandalized candidate more votes in intra-party than in inter-party contests, for *any* scandal severity, as long as voters place some value on sharing a candidate's party. Second, the amplification does not require that the voter's own party be at stake: it appears even when both candidates belong to the party the voter opposes, because the party label cancels from the comparison regardless of the voter's allegiance. This second prediction separates our substitutability account from one in which voters discipline only their own side (Dulay and Lee 2026).

The mechanism is that sanctioning is a choice, not a verdict: a scandal moves a vote only

if it is large enough to overturn the voter's prior preference between the two candidates on offer. Voters who lean firmly toward one side are unmoved, so what matters is how many sit near indifference. Across party lines, partisanship pulls most voters toward their own side and few remain near indifference. Within a party, the shared label removes the between-party utility gap, leaving faction, valence, and other attributes to decide; many more voters are nearly indifferent, and the same scandal readily tips them.

We test these implications with a pre-registered conjoint experiment on 3,000 U.S. adults, in which candidate party, faction, gender, scandal, and positive accomplishments vary independently across the two profiles in each task. Because party is randomized within the pair, roughly half of all contests place two same-party candidates in competition. This composition is itself randomized rather than something respondents select into, so whether a given task turns out to be an inter-party or an intra-party contest is unrelated to respondent or candidate characteristics, and comparisons across contest types are causal.

The results align with the model's central predictions. Attaching a scandal to a candidate lowers the probability that the candidate is chosen by 0.19 in inter-party contests and by 0.29 in intra-party contests—about 57% more punishment inside the party. In tasks where exactly one candidate is scandalized, that candidate loses 68% of the time when the alternative comes from the other party, 76% when both candidates come from the party the respondent opposes, and 80% when both come from the respondent's own party. The out-group result is the sharpest test of the mechanism: amplification produced by voters disciplining their own side should vanish there, while amplification produced by substitutability should persist. It persists.

Having established the causal inter/intra difference with the randomized design, we exploit the close connection between the experiment and the theoretical model to add a structural calibration of key model quantities. Because the model expresses party attachment, faction, and the scandal cost in the same utility units, the conjoint choices can be used to estimate those quantities directly, rather than only to test the qualitative pattern they imply.

A choice-level logit prices the average scandal at 81% of the utility of sharing a candidate's party—enough to decide a contest within a party, but not enough, for the average partisan, to justify crossing party lines. Pricing the scandals one at a time places them on either side of that threshold: the milder ones (falsified credentials, a violent protest while underage) cost clearly less than a party match, illegal donations cost about as much, and only sexual misconduct costs more.

The paper makes three contributions. Substantively, it shifts attention from the scandalized candidate to the choice set in which sanctioning takes place: credible co-partisan alternatives relax the constraint that punishment requires helping the opposing party. Theoretically, we offer a simple explanation for a pattern already documented observationally in the U.S.—that intra-party contests can discipline politicians more effectively than inter-party ones (Hirano and Snyder 2012). Methodologically, we treat the composition of the choice set as the treatment rather than as background. Conjoint designs ordinarily estimate how much each candidate attribute is worth (Schwarz and Coppock 2022); randomizing party across *both* profiles instead generates exogenous variation in the choice set, turning the design into a tool for studying when an attribute becomes electorally consequential—the context-dependence observational research cannot isolate.

Taken together, the results suggest a different way of thinking about the electoral check on misconduct. The recurring finding that voters barely punish wrongdoing has been read as a fact about voters—their information, attention, or willingness to hold their own side to account. We locate a substantial part of it in the ballots on which punishment has been measured: the same voters who will not cross party lines to remove a scandalized politician remove one readily when the replacement is a co-partisan. The binding constraint is not tolerance for misconduct but the price of acting on disapproval.

2 Related Literature

Political scandals involve transgressions of accepted norms or values that become publicly known, give rise to disapproval, and can damage the reputation of the political actors involved (Thompson 2013). Their public revelation gives voters negative information about candidates and an opportunity to sanction them at the ballot box, yet the recurring finding is that punishment is muted (Eggers 2014; Incerti 2020), with partisanship the usual explanation (Taber and Lodge 2006; Anduiza, Gallego and Muñoz 2013; Klar and McCoy 2021; Funck and McCabe 2022; Lee et al. 2023). This literature explains how a voter reacts to a scandalized candidate of a given party, but it has said little about the candidate on the other side of the ballot. Our premise is comparative: whether a voter acts on her disapproval depends on what she must accept instead.

The closest antecedent is Graham and Svolik (2020). Their concern is how much partisanship weakens the electoral check; ours is which choice sets leave it intact. Two of their findings set up our question: punishment of democratic violations is weak on average and few voters will cross party lines to impose it, but it is stronger when the two candidates share a party—a difference they read cautiously, because within same-party contests punishment follows a *partisan double standard* (each side judges the other party's violators more harshly than its own), so the voters most willing to punish a party's candidate are typically those who cannot vote in its primary. We take up the unexplained difference and argue that it comes from the choice set: a shared party label removes the partisan price of punishment for every voter, whatever her allegiance, which implies that the amplification should appear even inside the party the voter opposes.

Theoretical research on the supply of scandal asks how misconduct becomes public and when damaging information is released (Dewan and Myatt 2007). We study its demand, and we draw on the agency tradition in which elections discipline politicians only when voters have a viable replacement (Fearon 1999), making the identity of that replacement the variable

of interest. Our key theoretical observation—that a valence shock changes the most votes among voters near indifference—is familiar. The model's contribution is to restate it in the parameters our experiment measures, so that a qualitative prediction (larger penalties inside the party) comes with a quantitative condition (an inequality between party attachment and the scandal's utility cost) that the same data can evaluate.

The empirical literature on corruption voting reaches a similar conclusion from a different direction: voters punish corrupt candidates more readily when a clean alternative is available and acceptable on other grounds (Breitenstein 2019; Agerberg 2020). Reviewing that literature, De Vries and Solaz (2017) single out the viability of alternatives as a central open question.[1] Our design takes that step. Because the composition of the choice set is randomized, the availability and the party of the clean alternative are exogenous, and viability is defined by the attribute that dominates candidate choice in the United States—the party label.

The design also connects to work on intra-party politics, which distinguishes factions from other forms of internal conflict and cohesion (Webb and Bale 2021; Bolleyer and Kölln 2025). Perceived division can damage party evaluations and vote choice (Greene and Haber 2015), inconsistent internal rhetoric can make party positions harder for voters to locate (Jung and Somer-Topcu 2022), and a leadership candidate's own campaign communication can reallocate support among co-partisans (Loewen and Rubenson 2011). Recent work qualifies this view, showing that factions can improve candidate selection and voter welfare through information and discipline (Dewan and Squintani 2016; Bolleyer and Kölln 2025). We document a benefit of internal competition that requires no change in elite behavior: co-partisan competitors make scandals electorally actionable.

[1] Related work at the party-system level finds that voters switch more readily among proximate parties when their brands are less differentiated (Gunderson 2026).

3 A Model of Electoral Sanctioning

We present a model of how the partisan composition of a contest determines the electoral penalty for scandal. This section develops the logic and states the central results; Appendix A contains the formal statements, proofs, and extensions.

Voter i chooses between two candidates, $j \in \{1, 2\}$. Candidate j may carry a scandal s_j and a positive-valence cue r_j. Voter i's utility from candidate j is

$$u_{ij} = a_{ij} + v(s_j, r_j) + \varepsilon_{ij}, \tag{1}$$

where a_{ij} is the candidate's *clean appeal* (everything the voter values about the candidate apart from the scandal–cue bundle), v is the valence component generated by that bundle, and ε_{ij} is an idiosyncratic taste shock. Clean appeal decomposes as

$$a_{ij} = \beta_i \, \mathbf{1}\{P_j = P_i\} + \gamma_i \, \mathbf{1}\{P_j = P_i,\, F_j = F_i\} + \widetilde{a}_{ij}, \tag{2}$$

where P_j and F_j are candidate j's party and within-party faction, $\beta_i \geq 0$ is the value voter i attaches to a co-partisan, $\gamma_i \geq 0$ the additional value of a co-factional co-partisan, and $\widetilde{a}_{ij}$ collects all remaining components (e.g., ideological proximity, gender). For Independent voters, $\beta_i = \gamma_i = 0$. The party and faction terms can be microfounded either spatially, with candidates located at party-faction positions on an ideological dimension, or expressively, with party identity generating in-group utility; the two readings imply the same contest comparisons below, so nothing hinges on the choice.[2]

[2] Appendix A presents the spatial microfoundation, in which β_i is the utility gap created by the distance between the parties' platforms and γ_i by within-party factional distance. The empirical analysis is consistent with both readings.

For a scandal s attached to a candidate with cue r, define the *scandal cost*

$$q_s(r) \equiv v(0, r) - v(s, r) > 0, \tag{3}$$

the utility the scandal destroys, which may depend on the scandal's type and on the accompanying positive cue. Let m_i denote candidate 1's deterministic margin over candidate 2 when candidate 1 is clean. We assume that the difference in taste shocks has a symmetric, single-peaked density with cdf Φ, so candidate 1 is chosen with probability $\Phi(m_i)$ when clean and $\Phi(m_i - q)$ when scandalized. The electoral penalty for the scandal is therefore

$$\mathcal{P}_i = \Phi(m_i) - \Phi(m_i - q). \tag{4}$$

The central mechanism is a property of this expression that we call *pivotality*.

Lemma 1 (Pivotality)**.** *The penalty $\Phi(m) - \Phi(m - q)$ is single-peaked in the clean margin m, maximized at $q/2$ (where the scandal moves the most votes) and vanishing as $|m|$ grows.*

Intuitively, a scandal moves the most votes when the scandalized candidate and the alternative are close enough that the misconduct decides between them. It moves few votes when the candidate begins far ahead—most supporters stay—and few when the candidate begins far behind—most voters were already lost. Everything that follows is an application of this observation to the question of where, in utility terms, different contests place voters.

Contest composition matters because the party term in (2) either cancels or does not. To see this, split the clean margin into a party component and everything else. Let z_i denote candidate 1's clean margin excluding the party term—the margin generated by the two candidates' factions, gender, cues, and idiosyncratic appeal (Appendix A gives the explicit expression). In a *same-party contest* both candidates carry the same party label, β_i is common to both and drops out, and the clean margin is simply $m_i = z_i$. In an *opposing-party contest* the party term does not cancel and shifts the margin by β_i toward whichever

candidate is the voter's co-partisan: $m_i = z_i + \beta_i$ if candidate 1 shares the voter's party and $m_i = z_i - \beta_i$ if candidate 2 does. The consequence is a claim about where the three kinds of contest place voters relative to the pivotal region of Lemma 1.

Each of the three margins carries its own penalty. Write $\mathcal{P}_i^{\text{same}}$ for the penalty in a same-party contest, $\mathcal{P}_i^{\text{co}}$ for the penalty on the voter's co-partisan in an inter-party contest, and $\mathcal{P}_i^{\text{out}}$ for the penalty on the out-partisan in an inter-party contest. A voter facing a pair knows which of the three she is in. Proposition 1 compares them.

Proposition 1. *For any scandal cost $q_s(r)$ and otherwise-comparable candidates ($z_i = 0$):*

(i) $\mathcal{P}_i^{\text{same}} > \mathcal{P}_i^{\text{co}}$ *if and only if* $\beta_i > q_s(r)$*;*

(ii) $\mathcal{P}_i^{\text{same}} > \mathcal{P}_i^{\text{out}}$ *for every* $\beta_i > 0$*.*

Intuitively, part (i) requires party attachment to outweigh the scandal. Consider a scandalized co-partisan. If crossing party lines is nearly costless (β_i small), that candidate sits near indifference even in an inter-party race, so moving the contest inside the party adds little. If the scandal is severe enough that $q_s(r) > \beta_i$, it can push the voter across the party line on its own, so it retains its bite in an inter-party contest and again has less to gain from a same-party setting. Only when $\beta_i > q_s(r)$ does the scandalized co-partisan survive an inter-party contest but lose a same-party one. Both quantities have counterparts in the conjoint choices, so Section 7 can evaluate the inequality directly. Part (ii) needs no condition: a scandalized out-partisan starts at $z_i - \beta_i$, below indifference, and the scandal pushes her further down, so moving the contest inside the party can only help the scandal bind.

From voter behavior to the experimental estimand

Proposition 1 predicts behavior of a voter facing one pair of candidates. In the experiment, we will use the following estimand:

$$\mathcal{A}_i \;\equiv\; \mathcal{P}_i^{\text{same}} \;-\; \tfrac{1}{2}\big(\mathcal{P}_i^{\text{co}} + \mathcal{P}_i^{\text{out}}\big), \tag{5}$$

and refer to it as "amplification." This estimand pools inter-party tasks, and because party is randomized independently across the two profiles, in half of those tasks the scandalized candidate is the respondent's co-partisan and in half the out-partisan. The pooled contrast therefore compares the same-party penalty with the average of the two inter-party penalties.[3] The next result shows that the intra-party penalty is always greater than the average of inter-party penalties.

Corollary 1. *In the symmetric benchmark ($z_i = 0$), $\mathcal{A}_i > 0$ for every $\beta_i > 0$ and every $q_s(r)$, and $\mathcal{A}_i$ is strictly increasing in β_i.*

The benchmark $z_i = 0$ makes the statement clean but is not needed for the result. Appendix A (Remark A.1) shows that amplification remains positive throughout the range of non-party margins over which the scandal is capable of reversing the voter's ranking—that is, wherever the misconduct is decisive for the voter at all.

Amplification is positive even when $\beta_i \leq q_s(r)$, because of an asymmetry between the two inter-party contests. Displacing the margin downward by β_i costs more penalty than displacing it upward by β_i gains, since the penalty function falls away faster on the far side of the pivotal region than it climbs on the near side, so averaging the two always leaves less punishment than the same-party case. Substantively: when a scandal is severe enough to cost a co-partisan more votes across party lines than within them, it is also severe enough that the corresponding out-partisan was already beyond saving, and the second effect dominates.

Corollary 2. *Conditional on z_i and $q_s(r)$, $\mathcal{P}_i^{\text{same}}$ is the same whether the two candidates belong to the voter's own party or to the party she opposes.*

[3]The equal weighting reflects the design rather than a requirement of the argument. It is easy to show that amplification remains positive under any weighting whenever $\beta_i > q_s(r)$, and more generally for any weight on the co-partisan case below a threshold strictly greater than one half—which matters only for carrying the argument to observational settings, where the incidence of scandal across the two sides need not be balanced.

Corollary 2 is what separates substitutability from an in-group discipline account. If amplification reflected voters policing their own side—betrayal, moral expectations of one's own team—it should be confined to contests within the voter's party. Under substitutability it appears whenever the party label is common to both options, because the label cancels for every voter regardless of allegiance.[4]

Two further implications follow. First, by the monotonicity in Corollary 1, amplification increases with party attachment and is absent for pure Independents, for whom the party label never enters the margin. Second, Appendix A (Remark A.2) shows that amplification is non-monotonic in severity: it vanishes both for a costless scandal and for a disqualifying one, which costs the candidate every voter who preferred her whatever the contest. Amplification is therefore largest at intermediate severity; where the turn occurs depends on the distribution of taste shocks.

We now state the two pre-registered hypotheses the model was designed to test.

Hypothesis 1 (Intra-party amplification)**.** Amplification is positive ($\mathcal{A} > 0$): scandals reduce vote support more strongly in intra-party contests than in inter-party contests.

Hypothesis 2 (Out-group scope)**.** The amplification extends to out-group intra-party contests: scandals are more costly when both candidates belong to the party the respondent opposes than in inter-party contests.

4 The Conjoint Experiment

We test the model with a conjoint experiment embedded in a survey of 3,000 U.S. adults, recruited with quotas on gender, age, and region and fielded in Spring 2026. Appendix B reports quota targets, achieved sample composition, and descriptive counts. Each respondent

[4]Voters may still attend more closely to their own party's contests, so we treat exact equality as a benchmark rather than a sharp prediction; but amplification concentrated *entirely* in own-party pairs would reject the substitutability account.

completed four forced-choice *tasks*. A task is one paired electoral decision: respondents saw two fictional candidates competing in an election in their state and chose the one they would vote for, then rated each candidate's competence and honesty on 0–10 scales. Each candidate description is a *profile*, so the data comprise 24,000 profile observations nested in 12,000 tasks. At the profile level the vote outcome equals one for the chosen candidate and zero for the other, so the two outcomes within a task sum to one.

We randomized gender, party, faction, scandal, and positive-valence information across profiles. Party was Democratic or Republican; Democratic candidates were associated with the moderate or progressive wing of the party, Republican candidates with the moderate or MAGA wing.

We deliberately vary the content of the scandal treatment rather than restricting misconduct to a single domain, so that we can assess whether the choice-set mechanism extends across substantively different forms of norm violation. Much of the electoral-accountability literature has focused on corruption or financial wrongdoing (e.g., Eggers 2014; Breitenstein 2019; Incerti 2020), while other work examines sexual misconduct in particular (e.g., Klar and McCoy 2021). Our scenarios span illegal political conduct (illegal campaign donations), dishonesty about personal qualifications (falsified credentials), sexual misconduct (unwanted sexually explicit messages), and past personal misconduct (participation in a violent riot while underage).

The positive-valence attribute took five levels: no positive information, workplace anti-harassment enforcement, fact-checking accuracy, prevention of layoffs during a local downturn, and a community-service award. Appendix B reports full wording, outcome coding, and descriptive counts.

The central treatment of interest is the partisan composition of the profile pair. Because party was randomized independently across the two profiles, 5,911 of the 12,000 tasks are *intra-party contests* (both profiles share a party label) and 6,089 are *inter-party contests*. Intra-party contests divide further by the respondent's relation to the pair. Table 1

summarizes the three categories used to test the two hypotheses.

Table 1: Partisan composition of the experimental tasks.

Contest type	Two profiles in the task
Inter-party	One Democratic and one Republican candidate
In-group intra-party	Two candidates from the respondent's own party
Out-group intra-party	Two candidates from the party the respondent opposes

Note: The main test of Hypothesis 1 combines the two intra-party categories. The three-category analysis tests Hypothesis 2. For this analysis, Independents are assigned to a Democratic or Republican party camp using the pre-treatment faction-closeness item; strict robustness checks retain only party identifiers.

Independent respondents enter the main inter-party/intra-party analysis, which depends only on the composition of the candidate pair; Table 1 notes how they are coded in analyses that require the respondent's relation to the pair. Vote choice is the primary outcome throughout. Appendix C.4 reports secondary analyses of the post-choice competence and honesty ratings; because these were elicited after respondents voted, we treat them as supplementary evaluative outcomes rather than as mediators of candidate choice.

The experiment isolates the choice-set component of intra-party competition. It does not reproduce all features of real primaries, nomination contests, or leadership elections: respondents face a fixed paired ballot rather than a campaign with turnout selection, elite endorsements, media dynamics, campaign spending, party rules, or an explicitly described subsequent general election. Forced choice therefore identifies conditional candidate choice within that ballot, not turnout, abstention, third-party defection, or the decision to participate in an out-party contest. The claim is not that intra-party institutions mechanically restore accountability, but the narrower one that offering a co-partisan alternative makes sanctioning possible without handing the office to the opposing party.

5 Scandals Matter: General Patterns of Candidate Choice

Figure 1 summarizes how every randomized candidate attribute relates to vote choice. It reports *marginal means*: the average probability that a profile carrying each attribute level

is chosen, averaging over the other randomized attributes and the opposing profile. Unlike an average marginal component effect, which is a difference relative to a reference level, a marginal mean is a level of support and needs no reference category (Hainmueller, Hopkins and Yamamoto 2014; Leeper, Hobolt and Tilley 2020). The unit is the candidate profile and uncertainty is clustered by respondent.

The strongest separation is partisan: Democratic respondents are substantially more likely to choose Democratic profiles, and Republican respondents Republican profiles. Faction labels matter within the opposing party in particular—Democrats prefer a moderate Republican to a MAGA Republican, while Republicans prefer a moderate Democrat to a progressive Democrat—but these differences remain smaller than the between-party divide. The prominence of party congruence foreshadows the paper's mechanism and echoes evidence that partisan alignment conditions how voters respond to misconduct (Anduiza, Gallego and Muñoz 2013; Funck and McCabe 2022). Average gender differences are small and run in opposite directions across the two respondent groups.

Scandal is the most consistently damaging non-party attribute. Every scandal level attracts less support than the no-scandal level among both Democratic and Republican respondents, with sexual misconduct and illegal donations producing the lowest marginal means. That these two sit at the bottom parallels evidence from congressional elections that corruption and sex scandals damage reelection prospects more than political scandals (Basinger 2013), and the uniform negative shift is consistent with meta-analytic evidence that scandal reduces electoral support (Incerti 2020). Our estimand is the narrower probability of choosing a candidate within the presented pair. The next section shows that the magnitude of this response depends not only on the scandalized candidate but also on the alternative the voter is offered.

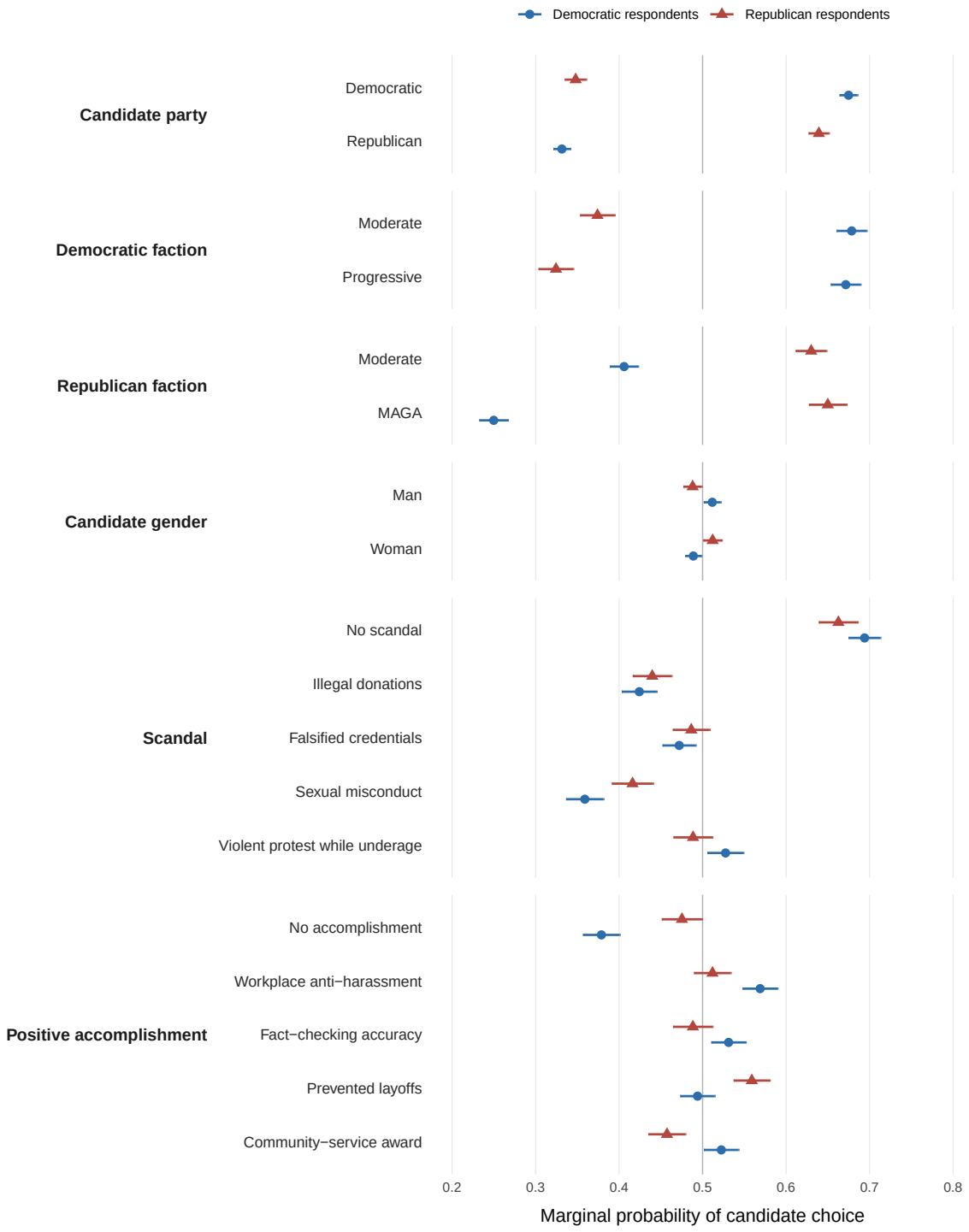


Figure 1: Vote-choice marginal means across candidate attributes, by respondent party.

Note: Points report profile-level probabilities of candidate choice for Democratic and Republican identifiers at each attribute level; horizontal bars are 95% confidence intervals with standard errors clustered by respondent. The vertical line marks 0.50. Democratic faction means are computed among Democratic candidate profiles and Republican faction means among Republican profiles. Democratic identifiers: N = 1,167 respondents and 9,336 profiles; Republican identifiers: N = 970 and 7,760.

6 The Mechanism: Substitutability and Amplification

The general patterns establish that respondents use the randomized candidate information and penalize all four scandals. We now turn to the paper's specific mechanism: candidate substitutability implies that the same negative information should change more votes when sanctioning does not require crossing party lines.

The main comparison uses the profile-level choice indicator and the linear probability model

$$Y_{ijt} = \alpha + \theta\, Scandal_{ijt} + \delta\, IntraParty_{it} + \tau\,(Scandal_{ijt} \times IntraParty_{it}) + X'_{ijt}\gamma + Z'_i\xi + \lambda_t + \varepsilon_{ijt}, \tag{6}$$

where i indexes respondents, $j \in \{1, 2\}$ profiles, and $t \in \{1, \ldots, 4\}$ task order. The outcome Y_{ijt} equals one when the profile is chosen. X contains the remaining randomized candidate attributes, Z respondent characteristics, and λ_t task-order fixed effects. Standard errors are clustered by respondent to account for repeated tasks and dependence between the two profiles in each task.

The corresponding profile-level average comparisons measure the change in a profile's predicted probability of being chosen when its scandal status changes, averaged over the observed distribution of the other variables (Hainmueller, Hopkins and Yamamoto 2014). In inter-party contests this penalty is θ; in intra-party contests it is $\theta + \tau$. Their difference, τ, is the amplification contrast that tests Hypothesis 1. In conjoint terminology, these quantities are conditional average marginal component effects (AMCEs), reported here as average comparisons. To test Hypothesis 2, we replace the binary contest indicator with the three categories in Table 1 and estimate a separate penalty in each. Appendix C reports specification robustness.

Figure 2 presents average predicted vote probabilities from Equation (6). Scandalized candidates fare worse in both settings, but the difference is visibly larger when the alternative

shares the candidate's party. In inter-party contests, the predicted probability of choosing a clean candidate is 0.66, against 0.46 for an otherwise comparable scandalized candidate; in intra-party contests the gap widens to 0.72 against 0.44.

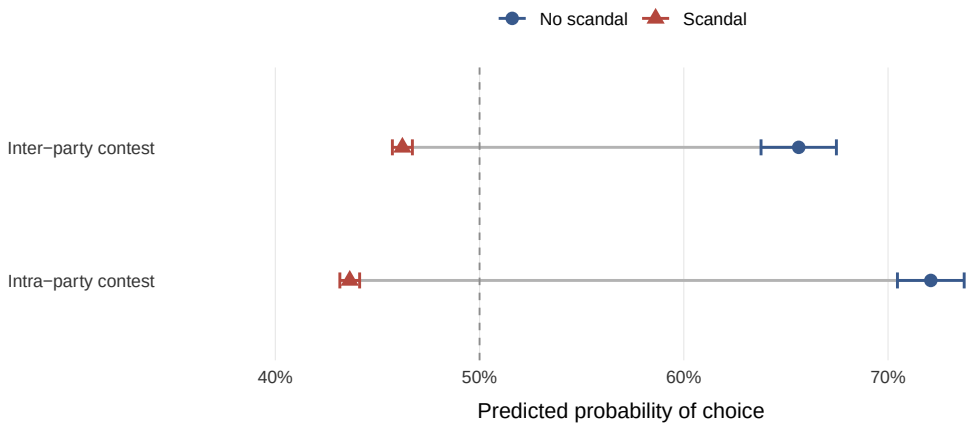


Figure 2: Average predicted vote probabilities by scandal presence and contest composition.

Note: Each point is the average profile-level prediction from Equation (6) after setting scandal status and contest composition to the displayed values and averaging over the observed values of the remaining variables. Gray lines connect the clean and scandalized predictions within each contest type, so their lengths show the absolute scandal penalty. Error bars are 95% confidence intervals; standard errors are clustered by respondent. N = 3,000 respondents; 12,000 tasks; 24,000 profile observations.

The vote penalty is −0.187 in inter-party contests and −0.292 in intra-party contests. Their difference—the intra-party penalty minus the inter-party penalty—is −0.106 (95% CI: −0.136 to −0.076). Negative values denote amplification because the intra-party penalty is more negative. In relative terms, the intra-party penalty is 56.7% larger than the inter-party penalty (95% CI: 34.9% to 78.6%). This is the pattern Corollary 1 predicts.

Positive accomplishments complement this result. If the mechanism operates by making candidate valence more consequential when party-based differentiation disappears, it should not be limited to negative information. Consistent with this implication, pooling the four positive cues, their effect increases from 7.9 percentage points in inter-party contests to 11.5 points in intra-party contests, a 3.6-point amplification (95% CI: 0.4 to 6.8; Table D6 in Appendix D.4).

Secondary ratings collected after vote choice show a parallel pooled pattern: on the 0–10

scales, the competence penalty is −1.47 in inter-party contests and −1.78 in intra-party contests, while the honesty penalty is −1.92 and −2.33, respectively. Appendix C.4 reports these estimates.

6.1 Out-group intra-party contests

The out-group comparison provides the strongest test of the substitutability mechanism, because it holds constant the respondent's relationship to the scandalized candidate and changes only the available alternative. Consider a scandalized candidate from the party the respondent opposes. In the relevant inter-party contest, rejecting that candidate means choosing a candidate from the respondent's own party; in an out-group intra-party contest, it means choosing another candidate from the same opposing party. What changes is only whether the two candidates are separated by the party boundary.

Corollary 2 and Hypothesis 2 distinguish the resulting predictions. An account operating only through voters' willingness to discipline their own party predicts no out-group amplification relative to the inter-party baseline. Under substitutability, amplification should persist: the shared out-party label cancels from the comparison, and the clean out-party candidate becomes a close substitute for the scandalized one.

Figure 3 reports the three profile-level average comparisons. The vote penalty is −0.187 in inter-party contests, −0.275 in out-group intra-party contests, and −0.310 in in-group intra-party contests. Both intra-party penalties are clearly distinguishable from the inter-party baseline: the out-group-minus-inter-party contrast is −0.089 (95% CI: −0.125 to −0.053) and the in-group-minus-inter-party contrast is −0.123 (−0.159 to −0.087). The 3.4-point difference between the two intra-party contexts is not distinguishable from zero (−0.034; 95% CI: −0.074 to 0.006; Table C5). The data are therefore consistent with the benchmark of Corollary 2—a same-party penalty that does not depend on whose party it is—while leaving room for a modest own-party surcharge, which we take up in Section 7. Restricting to Democratic and Republican identifiers gives a similar pattern: −0.163 (inter-party), −0.271

(out-group), and −0.309 (in-group). Excluding Independents widens the inter/intra separation, as the model predicts, because they are the respondents for whom contest composition creates the smallest partisan wedge.

The four-context decomposition among Democratic and Republican identifiers makes that comparison explicit (Table 2). For a scandalized out-party candidate, the penalty rises from −0.169 when the opposing profile is the respondent's co-partisan to −0.274 when it belongs to the same out-party. The own-party comparison is parallel: the penalty on a scandalized co-partisan rises from −0.144 when the alternative is an out-partisan to −0.313 when it is another co-partisan. A stable partisan double standard cannot by itself account for either within-relation increase, whereas removing the party cost of the alternative can. The comparison therefore locates much of the amplification in the composition of the choice set rather than in a change in how respondents evaluate the scandalized side.

This result sharpens the comparison with Graham and Svolik (2020), who read primaries as a limited democratic check because the double standard means the voters most willing to punish a party's candidate are typically barred from its primary (Graham and Svolik 2020, Figure 5). But the double standard concerns who the scandalized candidate is relative to the respondent; substitutability concerns who the available alternative is. The four-context comparison varies the latter while holding the former fixed.

Table 2: Four-context mechanism check among partisan identifiers.

Context	Penalty	95% CI low	95% CI high
Inter-party: scandalized candidate is respondent's party	−0.144	−0.166	−0.121
Inter-party: scandalized candidate is opposing party	−0.169	−0.203	−0.134
Intra-party: both candidates are respondent's party	−0.313	−0.346	−0.279
Intra-party: both candidates are opposing party	−0.274	−0.308	−0.240

Note: Entries are profile-level vote penalties for scandal among Democratic and Republican identifiers. Models use the main candidate controls, respondent controls, task-order fixed effects, and respondent-clustered standard errors.

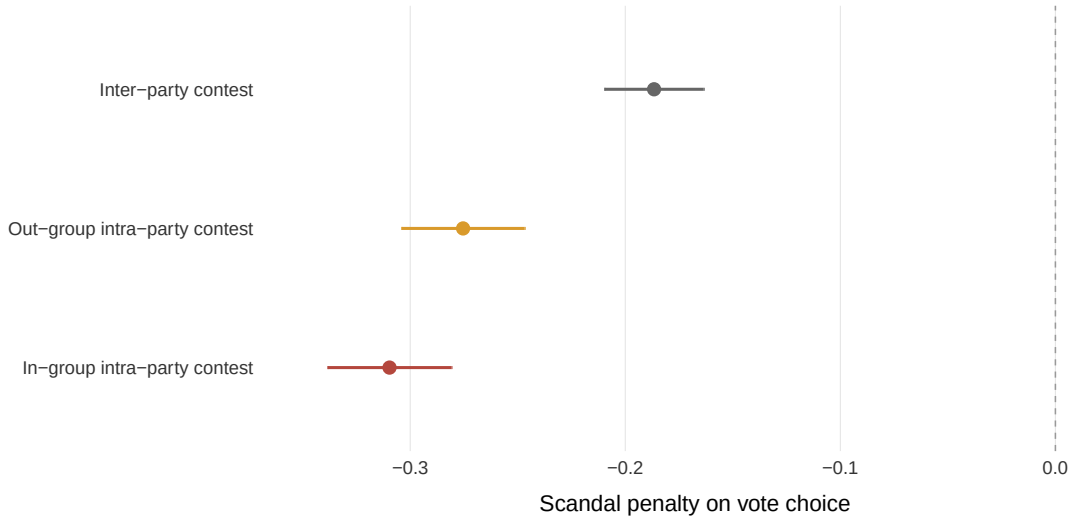


Figure 3: Vote penalty by respondent relation to the candidate pair.

Note: Each point is a profile-level average comparison: the change in the predicted probability of choosing a candidate when the profile moves from no scandal to scandal, estimated within the indicated respondent-pair context and averaged over the remaining variables. Error bars are 95% confidence intervals; standard errors are clustered by respondent. Independents are assigned to a party camp using the pre-treatment faction-closeness item. N = 3,000 respondents; 12,000 tasks; 24,000 profile observations.

6.2 How often does the scandalized candidate lose?

The estimates so far treat the candidate profile as the unit of analysis. The theory's basic unit, however, is the contest: one scandalized candidate facing one clean alternative. We restrict attention to the 4,172 tasks in which exactly one of the two profiles carries a scandal and ask a single question: how often does the scandalized candidate lose?

For each such task, let

$$L_{it} = \mathbf{1}\{\text{respondent } i \text{ does not choose the scandalized profile in task } t\}, \tag{7}$$

and for each of the three contest types in Table 1, let π_c denote the probability that the scandalized candidate loses among tasks of type c:

$$\pi_c = \Pr\big(L_{it} = 1 \mid \text{task } t \text{ is of type } c\big), \qquad c \in \{\text{inter-party, out-group intra, in-group intra}\}. \tag{8}$$

In the model, the same-party minus inter-party difference in loss probabilities *equals* the amplification $\mathcal{A}_i$ of Equation (5). Corollary 1 therefore predicts $\pi_{\text{intra}} > \pi_{\text{inter}}$, and Corollary 2 predicts that the out-group and in-group values sit near one another.

The data return exactly this pattern. The scandalized candidate loses 67.9% of the time in inter-party pairs, 75.6% in out-group intra-party pairs, and 80.0% in in-group intra-party pairs (Figure 4). A candidate who carries the only scandal in the pair usually loses everywhere; what contest composition changes is how reliably. Moving the same misconduct from an inter-party to a same-party pair raises the probability of defeat by eight to twelve percentage points—enough, in a competitive race, to be the margin between surviving a scandal and being removed.

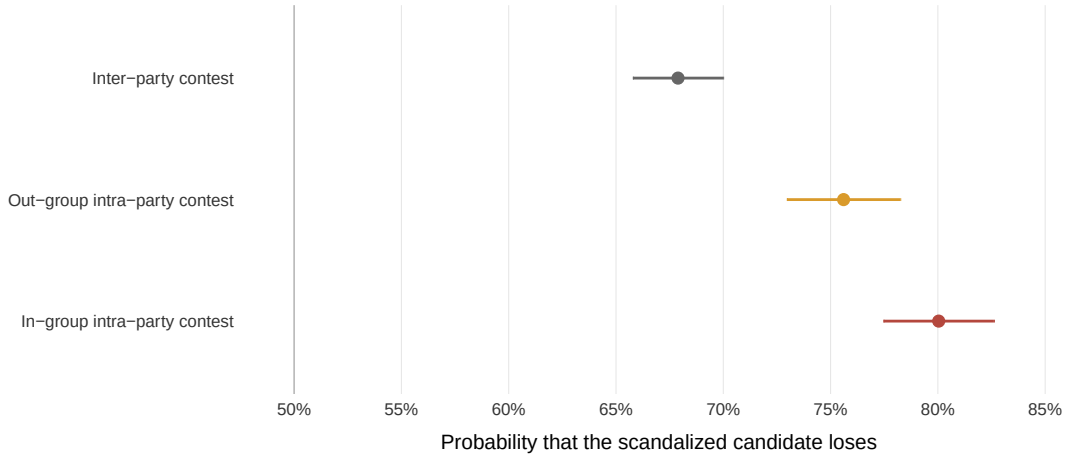


Figure 4: Task-level probability that the scandalized candidate loses.

Note: The sample is restricted to tasks with exactly one scandalized profile. Each point estimates π_c from Equation (8); bars are 95% confidence intervals with standard errors clustered by respondent. Inter-party: N = 1,651 respondents and 2,071 tasks; out-group intra-party: N = 977 and 1,074; in-group intra-party: N = 914 and 1,027. Respondent counts are not mutually exclusive across contexts.

The forced-choice format makes the out-group result demanding in another respect. Respondents could not abstain, reject both candidates, or choose their own party. Under those conditions a strategically motivated respondent might select the scandalized candidate to damage the opposing party or help it nominate a weaker general-election candidate—negative strategic voting of the kind documented in U.S. presidential primaries (Hillygus

and Treul 2014)—which could attenuate the penalty, eliminate it, or even reward the scandalized out-party candidate. The observed result runs the other way: required to choose between two out-party candidates, respondents rank the clean one more highly.

6.3 Heterogeneity and scope conditions

The model implies that amplification should be strongest among voters for whom party attachment creates the largest inter-party wedge. Among Democratic and Republican identifiers, amplification is small in the lowest partisan-strength tercile, with a 95% interval that narrowly includes zero (−0.059); it is larger in the middle tercile (−0.118) and largest in the highest (−0.204) (Figure 5). The high-strength estimate is distinguishable from the low-strength estimate, while adjacent tercile differences are less sharply estimated (Appendix E). We therefore treat the pattern as mechanism-consistent heterogeneity within the experiment rather than as evidence about participation or outcomes in actual primary electorates.

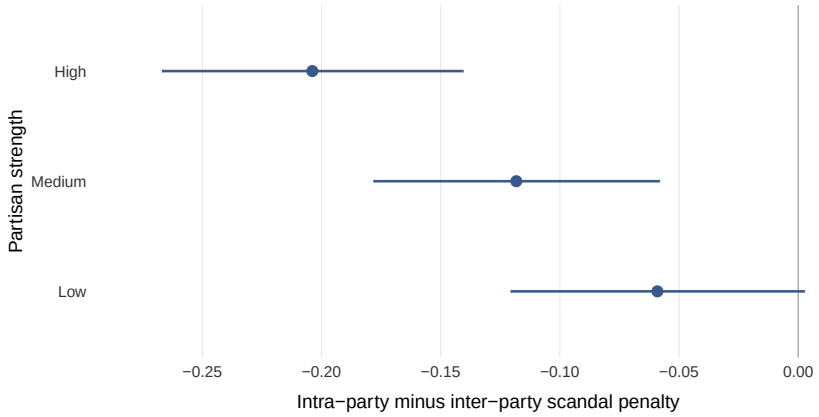


Figure 5: Intra-party amplification by partisan-strength tercile.

Note: Estimand: intra-party minus inter-party scandal penalty among Democratic and Republican identifiers. Negative values indicate larger scandal penalties in intra-party contests. Estimates come from a profile-level vote model interacting scandal, contest type, and partisan-strength tercile, with candidate controls, respondent controls, task-order fixed effects, and respondent-clustered standard errors. Error bars are 95% confidence intervals.

A second scope condition concerns what "same party" actually buys the voter. Removing the party label removes the largest single source of distance between two candidates, but

it does not make them interchangeable. Figure 1 shows that factional labels continue to differentiate candidate appeal even where party cannot, and because factions are randomized independently of party, we can hold fixed who is scandalized and vary only how close the available alternative stands to the voter—a within-party analogue of the contest-composition treatment.

Factional composition does condition the penalty, but its sign depends on whose faction is tainted. When the scandalized candidate comes from the respondent's own faction, offering an alternative from that same faction rather than from a rival one raises the penalty by 7.8 percentage points: the voter can withdraw support without also switching camps. When the scandalized candidate comes from a rival faction, the same substitution lowers the penalty by 6.3 points—consistent with the indifference logic of the model, since a voter already pulled firmly toward her own faction's candidate leaves the scandal little left to accomplish. The two contrasts differ by 14.1 percentage points (95% CI: 2.1 to 26.1). Their components are individually imprecise, so we read this as exploratory evidence that the factional configuration of the choice set can matter, rather than as a general or monotonic effect of factional proximity. Appendix E.4 reports the full four-configuration estimates and the task-level paired-loss complement.

6.4 Variation across scandal types

The pooled scandal indicator provides the main test of the contest-composition argument. Disaggregating the four randomized scenarios described above serves two narrower purposes: it shows whether the result depends on one particular form of misconduct, and it supplies scandal-specific costs for the threshold comparison in the structural analysis that follows.

Figure 6 reports a separate profile-level average comparison for each scenario relative to no scandal, within each contest composition. All four scenarios reduce the probability that a profile is chosen in both settings, and every estimate is more negative in intra-party than in inter-party contests. For falsified credentials, the penalty grows from -0.164 in inter-party

contests to −0.254 in intra-party contests; for illegal donations, from −0.199 to −0.330; for sexual misconduct, from −0.254 to −0.379; and for violent protest, from −0.127 to −0.220.

Sexual misconduct is the most damaging scenario in both contest types, followed by illegal donations; falsified credentials are intermediate, and violent protest—conduct that occurred while the candidate was underage, and so temporally distant and potentially more excusable—is least damaging. The amplification contrast is negative for every scenario (Appendix D), so the pooled result is not driven by any single form of misconduct. We do not compare amplification across scandal types, since with four types those pairwise differences are not directly estimated. One feature is nonetheless worth noting: amplification does not grow monotonically with severity—it is largest for illegal donations rather than for sexual misconduct, the most damaging scandal. By Remark A.2, this is expected: misconduct costly enough to rival the value of the party label moves voters across party lines even in inter-party contests, so it has less to gain from the label's removal. Appendix D reports the type-specific contrasts, paired-loss probabilities, and a two-scandal severity ranking; Appendix D.4 reports the corresponding vote-choice analyses for positive accomplishments.

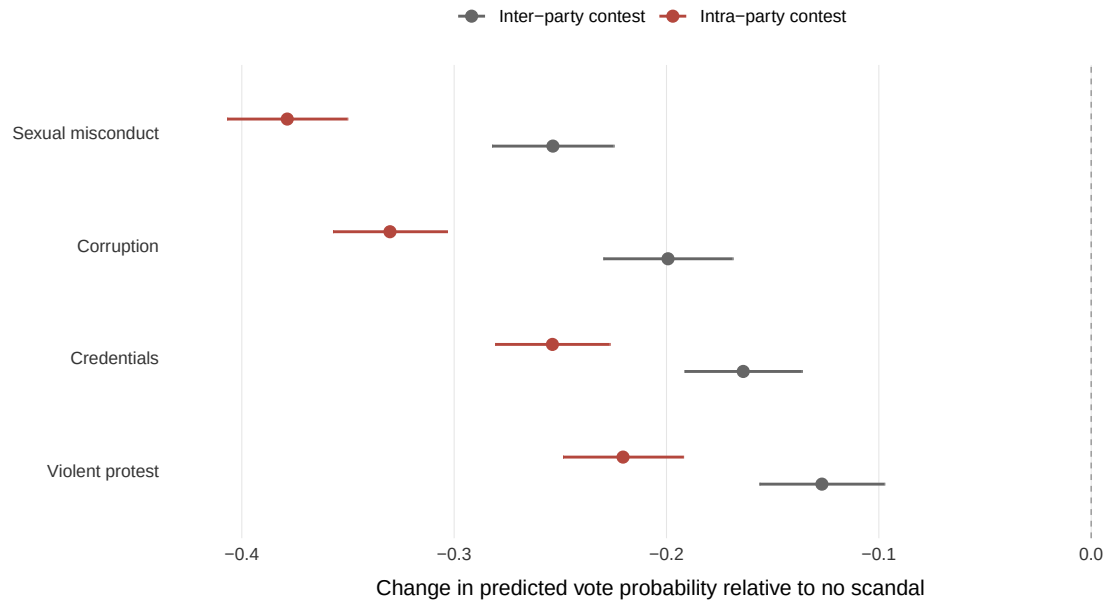


Figure 6: Vote penalties by scandal type and contest composition.

Note: Each point is a profile-level average comparison: the change in predicted vote probability from no scandal to the indicated scandal type, estimated within contest composition and averaged over the remaining variables. Error bars are 95% confidence intervals; standard errors are clustered by respondent. N = 3,000 respondents; 12,000 tasks; 24,000 profile observations.

7 Structural Calibration of the Mechanism

The design-based analyses already establish the paper's causal empirical result: how the scandal penalty changes across randomly assigned contest types. We now add a complementary structural analysis addressing two further questions. First, is the utility cost of a scandal large enough to offset the advantage of sharing a candidate's party? Second, is the mechanism quantitatively sufficient? In the model a scandal destroys the same utility in every contest, and the intra-party penalty is larger only because the shared label places voters closer to indifference. The estimated version preserves that restriction, so comparing the predicted gap with the observed one asks whether the mechanism, on its own, accounts for the amplification we measure.[5]

[5]The approach follows related candidate-choice work that prices valence against partisan considerations (Graham and Svolik 2020) and a growing literature that combines conjoint experiments with structural

We first estimate the pooled utility weights and express the scandal cost relative to the other attributes, then translate those weights into predicted choice probabilities and compare them with the reduced-form penalties. Appendix F additionally allows the scandal cost to vary across types and compares each type-specific cost with the common party-alignment weight.

The calibration uses the same experimental choices as the reduced-form analyses but organizes them at the level at which respondents decide. Each paired task contributes one binary observation rather than two profile rows. Because the choice identifies the relative utility of the two candidates, each attribute enters as the first profile's value minus the second profile's value; attributes shared by both profiles therefore cancel.

When the difference in unobserved utility between the two profiles is logistic, the probability that respondent i chooses the first profile in task t is

$$\begin{aligned}\Pr(y_{it}=1) = \Lambda\big[\,&\alpha + \beta\,\Delta\text{PartyMatch}_{it} + \gamma\,\Delta\text{FactionMatch}_{it} \\ &+ \kappa\,\Delta\text{Scandal}_{it} + \rho\,\Delta\text{PositiveValence}_{it}\big].\end{aligned} \tag{9}$$

Here Λ is the logistic cdf. For example, $\Delta\text{Scandal} = 1$ when only the first profile is scandalized, -1 when only the second is, and zero when their scandal status matches. We estimate Equation (9) by maximum likelihood and cluster standard errors by respondent to account for the four choices made by each participant. The equivalent candidate-long conditional logit produces the same estimates (Appendix F).

The coefficients are pooled utility weights: β is the utility of a co-partisan candidate, γ the additional utility of a co-factional co-partisan, $q = -\kappa$ the pooled scandal cost, and ρ the value of a positive-valence cue.[6] A positive coefficient raises the probability that the first profile is chosen; because scandal lowers utility, κ is negative and its cost q is reported as positive. These coefficients are expressed on the normalized logit-utility scale. We report

models (Acharya, Hainmueller and Xu 2026).

[6]Coefficients pool respondents and attributes (scandals and positive cues).

them as marginal rates of substitution (MRS)—the utility cost of a scandal expressed in units of another candidate attribute. For example, $q/\beta = 0.81$ means the average scandal removes utility equal to 81% of the advantage conferred by a shared party label. It does not mean scandal lowers vote probability by 81 percentage points; that change also depends on both candidates' attributes and is largest when the pair is closely matched. Party and faction match are coded as zero for Independents, and the main estimates restrict to Democratic and Republican identifiers so that β and q describe the same population. Results are nearly identical in the full sample and under alternative treatments of Independents (Appendix F).

Equation (9) is a standard random-utility, conditional-logit representation of paired choice when the difference in unobserved utility is logistic (McFadden 1974; Train 2009). Design-based conjoint analysis does not require this representation (Hainmueller, Hopkins and Yamamoto 2014).

Table 3 reports the calibration. A co-partisan candidate is worth $\beta = 1.45$ in normalized utility terms, a co-factional co-partisan adds $\gamma = 0.34$, and the average scandal costs $q = 1.17$. The average scandal cost is 0.81 times the utility of sharing a candidate's party (95% CI: 0.73 to 0.89)—not enough, for the average partisan, to offset the party advantage—and 0.65 times the combined value of sharing both party and faction. Measured against the differences that remain *inside* a party, however, the same scandal is large: it costs roughly 3.4 times the additional utility of a co-factional candidate.

Because the calibration places every attribute on one scale, it also lets us price misconduct against the good news a candidate can offer. For the treatments randomized here, the average scandal carries 2.64 times the utility weight of the average positive accomplishment (95% CI: 2.08 to 3.19)—bad news is worth about two and a half times the good news candidates ordinarily advertise. The two cues appear to work largely independently. Attaching a positive accomplishment to an otherwise scandalized profile recovers support in both contest types, but it does not measurably attenuate the marginal scandal penalty; the joint effects are consistent with approximate additivity, and the utility the positive cue restores falls short

of what the scandal destroys (Panel B of Table D7 in Appendix D.4).

Read against the long-standing conjecture that voters will forgive misconduct in exchange for valued outcomes (Rundquist, Strom and Peters 1977), our exploratory estimates suggest that they discount it, but not enough to save the candidate.[7] One clarification follows from these magnitudes. The choice-set mechanism amplifies valence in both directions—positive cues, too, count for more once the party label cancels (Section 6)—so the fact that amplification is larger in absolute terms for scandal reflects the larger utility shock the scandal treatments deliver, not a punishment-specific asymmetry in how voters process information. A scandal small relative to the party divide but large relative to within-party differences is exactly the configuration that produces intra-party amplification.

The type-specific estimates in Appendix F.3 compare each scandal cost with party-match utility to assess which scandal types cross the threshold at which misconduct outweighs partisan alignment. Falsified credentials and violent protest fall clearly short of it, illegal donations sit close to it, and only sexual misconduct exceeds it on the point estimate.

Table 3: Utility-weight estimates and implied marginal rates of substitution.

Utility component	Estimate	Clustered SE	95% CI
Co-partisan utility, β	1.449	0.054	[1.343, 1.556]
Additional co-faction utility, γ	0.343	0.053	[0.239, 0.446]
Scandal utility, κ	-1.172	0.050	$[-1.269, -1.075]$
Positive-valence cue utility, ρ	0.445	0.046	[0.356, 0.534]
MRS: scandal / party, $\lvert\kappa/\beta\rvert$	0.809	0.041	[0.729, 0.888]
MRS: scandal / total alignment, $\lvert\kappa/(\beta+\gamma)\rvert$	0.654	0.031	[0.593, 0.715]
MRS: scandal / faction increment, $\lvert\kappa/\gamma\rvert$	3.419	0.545	[2.352, 4.487]
MRS: scandal / positive-valence cue, $\lvert\kappa/\rho\rvert$	2.635	0.283	[2.080, 3.189]

Note: Choice-level logit estimates of equation (9), with covariates constructed as differences between the two profiles in each task. N = 2,137 respondents and 8,548 choice tasks. Standard errors clustered by respondent.

As a related substantive benchmark, Graham and Svolik (2020) estimate that a candidate who violates democratic principles loses about 11.7% of the vote overall, and that only

[7] Recent experimental work similarly examines trade-offs between positive substantive performance and procedural transparency in evaluations of party trust (Kölln 2026).

13.1% of respondents are willing to cross party lines to punish a co-partisan violator. Their experiment concerns democratic-principle violations, but both sets of estimates reveal the same constraint: a negative valence signal can matter substantially while remaining too small to overcome the partisan advantage in many inter-party choices.

7.1 Model-implied amplification

The calibration also provides a quantitative benchmark for the observed penalties. To match the population underlying the reduced-form penalties, this comparison uses the full-sample strict-coding specification, in which party and faction matches are set to zero for Independents; Appendix F reports the corresponding partisan-only and alternative-coding results. For each task, we predict choice with the profile clean and scandalized, holding the remaining profile attributes and the opposing profile fixed, and average the difference within each contest context. Any predicted inter/intra gap must therefore come from the alignment utilities and the logit link alone: the model cannot fit amplification by assigning scandal a larger utility cost inside the party. Table 4 compares these model-implied quantities with the reduced-form estimates from Section 6.1.

Table 4: Model-implied and observed scandal penalties by contest composition.

Contest context	Predicted	Observed	Difference
Inter-party	−0.192	−0.187	−0.005
Out-group intra-party	−0.259	−0.275	0.016
In-group intra-party	−0.257	−0.310	0.052

Note: Strict coding, full sample. Predicted penalties simulate a clean-to-scandal change for one profile, preserving the observed distribution of alignment and positive-valence attributes. The model contains no scandal-by-contest interaction: differences across contests arise only from the estimated alignment utilities passing through the logit choice probability. Observed values are the adjusted reduced-form vote penalties from Section 6.1.

Despite the restrictive specification, the model predicts a 19-point inter-party penalty, a 26-point out-group intra-party penalty, and a 26-point in-group intra-party penalty. The corresponding observed penalties are 19, 28, and 31 points. Thus the estimated utility weights account for most of the inter/intra gap, especially the out-group amplification that

is hardest for an in-group-discipline account to explain.

The model predicts nearly identical out-group and in-group intra-party penalties, while the observed in-group penalty is about five points larger. This residual shows that the parsimonious benchmark does not capture every process operating in respondents' own-party contests. It may reflect greater attention to one's own party, identity-based betrayal, stronger moral expectations of one's own side, or context-specific error variance.[8]

8 Discussion and Conclusion

Why do voters so rarely remove politicians who misbehave? Much of the answer, we have argued, lies in the ballots on which disapproval must be expressed rather than in voters' tolerance for misconduct. A scandal changes an election only if it can tip voters near indifference, and a shared party label is what places them there. A pre-registered conjoint experiment on 3,000 Americans confirms this conjecture: the vote penalty for scandal is roughly 57% larger when the alternative is a co-partisan, and a scandalized candidate facing a clean opponent loses 68% of the time across party lines against 76 to 80% within a party. The amplification holds even when both candidates come from the party the respondent opposes.

Three implications follow. First, nearly all existing estimates of scandal punishment come from settings in which punishing means handing office to the other party—precisely the choice sets where the electoral check is weakest. That voters barely sanction misconduct (Eggers 2014; Incerti 2020) may describe the ballots more than the electorate. Second, polarization does not reduce voters' distaste for misconduct; it raises the price of acting on it, and only across the party line. As attachment grows, the inter-party penalty falls toward zero while the intra-party penalty is untouched: polarization relocates accountability rather than disabling it. Third, the supply of alternatives is institutional. Whether a scandalized

[8] Appendix F reports a heteroskedastic-logit check allowing the error scale to vary by contest context. The check predicts a larger in-group intra-party penalty and reduces most of the residual, suggesting that part of the extra in-group penalty may be a scale artifact rather than a distinct scandal cost.

incumbent faces a credible co-partisan challenger depends on nomination rules, ballot access, and how much internal competition parties tolerate. Internal division may damage a party's standing, but factions and challengers also populate ballots with substitutes—and substitutes are what make disapproval actionable.

Co-partisanship, however, is a proxy for the mechanism and not the mechanism itself. The U.S. same-party contest is a clean setting in which to isolate it, because when the party-label component cancels the largest source of distance in the clean utility margin disappears. But a co-partisan need not be an equally close substitute—ideological and factional differences within a party can keep replacement costly, which is why our factional estimates describe a scope condition rather than a uniform within-party effect. By the same logic, candidates from different parties can be close substitutes in multi-party systems. Moderates facing proximate alternatives should find sanctioning cheap; ideologically extreme voters should find it difficult. Electoral institutions, in turn, structure which substitutes are available.

That last contrast points to where we believe the argument should be taken next. If accountability depends on the menu, institutional variation in menus is the natural laboratory. For instance, two-round systems imply a reversal within a single election: the same misconduct should be punished in the first round, where a voter can defect to a proximate ally, and tolerated in the runoff, where defection once again means crossing the party line. Single-transferable-vote and open-list contests, in which co-partisans routinely run against one another, make intra-party sanctioning the ordinary case rather than the exception, as do sub-national elections in fragmented multiparty systems.

References

Acharya, Avidit, Jens Hainmueller and Yiqing Xu. 2026. "Learning Preferences from Conjoint Data: A Hybrid Structural Deep Learning Approach.". arXiv preprint arXiv:2604.10845. **URL:** *https://arxiv.org/abs/2604.10845*

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

Anduiza, Eva, Aina Gallego and Jordi Muñoz. 2013. "Turning a Blind Eye: Experimental Evidence of Partisan Bias in Attitudes toward Corruption." *Comparative Political Studies* 46(12):1664–1692.

Basinger, Scott J. 2013. "Scandals and Congressional Elections in the Post-Watergate Era." *Political Research Quarterly* 66(2):385–398.

Bolleyer, Nicole and Ann-Kristin Kölln. 2025. "The Study of Intraparty Frictions: Conceptual Reflections on Preference Heterogeneity, Disagreement, and Conflict." *Perspectives on Politics* 23(3):885–900.

Breitenstein, Sofia. 2019. "Choosing the Crook: A Conjoint Experiment on Voting for Corrupt Politicians." *Research & Politics* 6(1):2053168019832230.

De Vries, Catherine E. and Hector Solaz. 2017. "The Electoral Consequences of Corruption." *Annual Review of Political Science* 20:391–408.

Dewan, Torun and David P. Myatt. 2007. "Scandal, Protection, and Recovery in the Cabinet." *American Political Science Review* 101(1):63–77.

Dewan, Torun and Francesco Squintani. 2016. "In Defense of Factions." *American Journal of Political Science* 60(4):860–881.

Dionne, Jr., E. J. 1987. "Gary Hart, the Elusive Front-Runner." The New York Times Magazine, May 3.

Dulay, Dean and Seulki Lee. 2026. "Partisanship versus Principles: A Study of Voter Response to Intra-Party Corruption Allegations." *Party Politics* 32(1):165–173.

Eggers, Andrew C. 2014. "Partisanship and Electoral Accountability: Evidence from the UK Expenses Scandal." *Quarterly Journal of Political Science* 9(4):441–472.

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Adam Przeworski, Susan C. Stokes and Bernard Manin. Cambridge University Press pp. 55–97.

Funck, Amy S. and Katherine T. McCabe. 2022. "Partisanship, Information, and the Conditional Effects of Scandal on Voting Decisions." *Political Behavior* 44(3):1389–1409.

Graham, Matthew H. and Milan W. Svolik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States." *American Political Science Review* 114(2):392–409.

Greene, Zoe and Matthias Haber. 2015. "The Consequences of Appearing Divided: An Analysis of Party Evaluations and Vote Choice." *Electoral Studies* 37:15–27.

Gunderson, Jacob R. 2026. "Party Brands, Issue Salience, and Electoral Volatility." *Journal of European Public Policy* 33(7):1950–1980.

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

Hillygus, D. Sunshine and Sarah A. Treul. 2014. "Assessing Strategic Voting in the 2008 US

Presidential Primaries: The Role of Electoral Context, Institutional Rules, and Negative Votes." *Public Choice* 161(3–4):517–536.

Hirano, Shigeo and James M. Snyder. 2012. "What Happens to Incumbents in Scandals?" *Quarterly Journal of Political Science* 7(4):447–456.

Incerti, Trevor. 2020. "Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design." *American Political Science Review* 114(3):761–774.

Jung, Jae-Hee and Zeynep Somer-Topcu. 2022. "United We Stand, Divided We Fall? The Effects of Parties' Brexit Rhetoric on Voters' Perceptions of Party Positions." *Journal of Elections, Public Opinion and Parties* 32(3):596–614.

Kirkland, Patricia A. and Alexander Coppock. 2018. "Candidate Choice Without Party Labels." *Political Behavior* 40(3):571–591.
URL: *https://doi.org/10.1007/s11109-017-9414-8*

Klar, Samara and Alexandra McCoy. 2021. "Partisan-Motivated Evaluations of Sexual Misconduct and the Mitigating Role of the #MeToo Movement." *American Journal of Political Science* 65(4):777–789.

Kölln, Ann-Kristin. 2026. "A Party's Pledge Fulfillment and Procedural Transparency Affect Voters' Trust." *West European Politics* pp. 1–26.

Lee, Amber Hye-Yon, Allison Harell, Laura B. Stephenson, Daniel Rubenson and Peter John Loewen. 2023. "Motivated to Forgive? Partisan Scandals and Party Supporters." *Political Psychology* 44(4):729–747.

Leeper, Thomas J., Sara B. Hobolt and James Tilley. 2020. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis* 28(2):207–221.

Loewen, Peter John and Daniel Rubenson. 2011. "For Want of a Nail: Negative Persuasion in a Party Leadership Race." *Party Politics* 17(1):45–65.

McFadden, Daniel. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*, ed. Paul Zarembka. New York: Academic Press pp. 105–142.

Rundquist, Barry S., Gerald S. Strom and John G. Peters. 1977. "Corrupt Politicians and Their Electoral Support: Some Experimental Observations." *American Political Science Review* 71(3):954–963.

Schwarz, Susanne and Alexander Coppock. 2022. "What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of Sixty-Seven Factorial Survey Experiments." *The Journal of Politics* 84(2):655–668.

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

Thompson, John B. 2013. *Political scandal: Power and visibility in the media age*. Polity Press.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. 2 ed. Cambridge: Cambridge University Press.

Webb, Paul and Tim Bale. 2021. Conflict and Cohesion within Parties. In *The Modern British Party System*. 2 ed. Oxford University Press pp. 205–236.

A The Model: Formal Statement and Proofs

This appendix states the model formally, proves the results in Section 3, and records extensions: the spatial microfoundation, the severity comparative static, and heterogeneity in party attachment.

A.1 Setup

Voters are indexed by i and choose between two candidates, $j \in \{1, 2\}$. Candidate j has a party P_j, a within-party faction F_j, a scandal type $s_j \in \mathcal{S}$, and a positive-valence cue $r_j \in \mathcal{R}$, where $s_j = 0$ denotes no scandal and $r_j = 0$ no cue. Voter i's utility from candidate j is

$$u_{ij} = a_{ij} + v(s_j, r_j) + \varepsilon_{ij}, \tag{A.1}$$

with $v(0, 0) = 0$. Positive cues raise v, scandals lower it, and the effect of a scandal may depend on the accompanying cue. Clean appeal is

$$a_{ij} = -\lambda \left| x_i - x_{P_j F_j} \right| + \beta_i \, \mathbf{1}\{P_j = P_i\} + \gamma_i \, \mathbf{1}\{P_j = P_i,\, F_j = F_i\}, \tag{A.2}$$

where x_i is voter i's ideological location, $x_{P_j F_j}$ the location associated with candidate j's party-faction label, $\lambda \geq 0$ the weight on ideological proximity, $\beta_i \geq 0$ the value of a co-partisan, and $\gamma_i \geq 0$ the additional value of a co-factional co-partisan. For Independent voters, the party and faction indicators are zero: $\beta_i = \gamma_i = 0$, and only the ideological term operates. Gender and other clean attributes can be added to a_{ij}; they do not affect the contest-composition logic and are suppressed.

Equation (A.2) nests two microfoundations. In the spatial reading, the four party-faction locations x_{PF} are ordered on the real line, between-party distances exceed within-party distances, and β_i can be set to zero: the party gap in clean appeal is generated by λ and the platform geometry. In the identity reading, λ is small and β_i, γ_i carry the alignment utility

directly. Both readings imply that a shared party label cancels from within-pair comparisons while an unshared one does not, which is the only property the results use.

Voter i chooses candidate 1 if and only if $u_{i1} \geq u_{i2}$. Let $\eta_i \equiv \varepsilon_{i2} - \varepsilon_{i1}$.

Assumption A.1. η_i has a continuous density ϕ that is symmetric about zero, strictly single-peaked, and has full support; write Φ for its cdf.

Assumption A.1 is satisfied by the logistic and normal cases. Two consequences are used repeatedly in what follows. First, symmetry about zero together with strict single-peakedness places the mode at zero, so ϕ is strictly decreasing in $|\cdot|$: whenever $|x| < |y|$ we have $\phi(x) > \phi(y)$. Second, full support is what makes that inequality strict everywhere rather than only on the interior of a smaller set. Candidate 1's choice probability is

$$V_i(s_1, s_2, r_1, r_2) = \Phi\big(a_{i1} - a_{i2} + v(s_1, r_1) - v(s_2, r_2)\big). \tag{A.3}$$

For a scandal type $s \neq 0$, define the *scandal cost* for a candidate carrying cue r as

$$q_s(r) \equiv v(0, r) - v(s, r) > 0. \tag{A.4}$$

Fix candidate 2's attributes and candidate 1's cue, and let

$$m_i \equiv a_{i1} - a_{i2} + v(0, r_1) - v(s_2, r_2) \tag{A.5}$$

be candidate 1's deterministic *clean margin*. The electoral penalty from assigning scandal s to candidate 1 is

$$\mathcal{P}_i(s, r_1; m_i) \equiv V_i(0, s_2, r_1, r_2) - V_i(s, s_2, r_1, r_2) = \Phi(m_i) - \Phi\big(m_i - q_s(r_1)\big). \tag{A.6}$$

To simplify notation, write $q \equiv q_s(r_1)$ and $g_q(m) \equiv \Phi(m) - \Phi(m - q)$.

A.2 Pivotality

Lemma 1 (Pivotality)**.** *For any $q > 0$, $g_q(m) = \int_{m-q}^{m} \phi(t)\,dt$ is strictly single-peaked and symmetric about $m = q/2$: it is increasing for $m < q/2$, decreasing for $m > q/2$, and $g_q(m) \to 0$ as $|m| \to \infty$.*

Proof. $g_q'(m) = \phi(m) - \phi(m-q)$. By Assumption A.1 ϕ is symmetric about zero and strictly single-peaked, hence strictly decreasing in $|\cdot|$, so $\phi(m) > \phi(m-q)$ if and only if $|m| < |m-q|$; squaring, this is $m^2 < m^2 - 2mq + q^2$, that is $2mq < q^2$, so, dividing by $q > 0$, it holds if and only if $m < q/2$. For symmetry, substitute $u = -t$ in

$$g_q(q/2 - x) = \int_{-x-q/2}^{-x+q/2} \phi(t)\,\mathrm{d}t = \int_{x-q/2}^{x+q/2} \phi(u)\,\mathrm{d}u = g_q(q/2 + x),$$

using $\phi(-u) = \phi(u)$. Hence $g_q(m) = G(|m - q/2|)$ for a strictly decreasing G. For the tail limits, $g_q(m) \le \int_{m-q}^{\infty} \phi = 1 - \Phi(m-q) \to 0$ as $m \to +\infty$, and $g_q(m) \le \int_{-\infty}^{m} \phi = \Phi(m) \to 0$ as $m \to -\infty$. □

Substantively, a scandal of fixed cost moves the most votes when the clean contest is near parity in utility terms. More precisely, g_q integrates ϕ over a window of fixed width q whose right endpoint is the clean margin; the area is greatest when that window is centered on the mode of ϕ, which occurs at $m = q/2$—the contest in which the scandal moves the candidate from $+q/2$ to $-q/2$, straddling indifference symmetrically. For small q this is close to the even contest, and the two coincide in the limit.

A.3 Contest composition

Write $a_{ij} = \widetilde{a}_{ij} + \beta_i \mathbf{1}\{P_j = P_i\}$, where $\widetilde{a}_{ij}$ collects ideological proximity, the factional term, and all non-party components. Define

$$z_i \equiv \widetilde{a}_{i1} - \widetilde{a}_{i2} + v(0, r_1) - v(s_2, r_2), \tag{A.7}$$

candidate 1's clean margin excluding the party term.

In a *same-party contest* ($P_1 = P_2$), the party term is common to both candidates and cancels, whether or not the shared party is the voter's:

$$m_i^{\text{same}} = z_i. \tag{A.8}$$

In an *opposing-party contest*, the party term shifts the margin toward the voter's co-partisan:

$$m_i^{\text{opp,co}} = z_i + \beta_i, \qquad m_i^{\text{opp,out}} = z_i - \beta_i, \tag{A.9}$$

according to whether candidate 1 is the voter's co-partisan or the out-partisan.

Lemma A.1 (Substitutability)**.** Fix a scandal cost $q > 0$. For any two contest types A and B with clean margins m_A and m_B, the scandal penalty is larger in contest A than in contest B if and only if $|m_A - q/2| < |m_B - q/2|$.

Proof. By Lemma 1, $g_q(m) = G(|m - q/2|)$ with G strictly decreasing, so $g_q(m_A) > g_q(m_B)$ if and only if $|m_A - q/2| < |m_B - q/2|$. □

Lemma A.1 is the general characterization: scandals matter most where the scandalized candidate and the alternative are close substitutes, in the precise sense that the clean margin lies near the penalty-maximizing point $q/2$.

A.4 Comparing contests

With $z_i = 0$ the three penalties are $\mathcal{P}_i^{\text{same}} = g_q(0)$, $\mathcal{P}_i^{\text{opp,co}} = g_q(\beta_i)$, and $\mathcal{P}_i^{\text{opp,out}} = g_q(-\beta_i)$.

Proposition 1. *In the symmetric benchmark $z_i = 0$, for any $q > 0$:*

1. $\mathcal{P}_i^{\text{same}} > \mathcal{P}_i^{\text{opp,co}}$ *if and only if* $\beta_i > q$*;*

2. $\mathcal{P}_i^{\text{same}} > \mathcal{P}_i^{\text{opp,out}}$ *for all* $\beta_i > 0$.

Proof. Part 1. $g_q(0) > g_q(\beta_i)$ if and only if $q/2 < |\beta_i - q/2|$. If $\beta_i > q$ then $\beta_i - q/2 > q/2 > 0$, so $|\beta_i - q/2| = \beta_i - q/2 > q/2$ and the inequality holds. If $0 < \beta_i \leq q$ then $\beta_i - q/2 \in (-q/2, q/2]$, so $|\beta_i - q/2| \leq q/2$ and it fails. Hence the strict inequality holds exactly when $\beta_i > q$.

Part 2. $|-\beta_i - q/2| = \beta_i + q/2 > q/2$ for every $\beta_i > 0$, so $g_q(0) > g_q(-\beta_i)$. □

A.5 The experimental estimand

An inter-party task realizes either $\mathcal{P}_i^{\text{opp,co}}$ or $\mathcal{P}_i^{\text{opp,out}}$, and pooling such tasks averages the two with weights fixed by the assignment mechanism. Because scandal is randomized independently of party, those weights are equal, and the estimand identified by the design is

$$\mathcal{A}_i \;\equiv\; \mathcal{P}_i^{\text{same}} - \tfrac{1}{2}\big(\mathcal{P}_i^{\text{opp,co}} + \mathcal{P}_i^{\text{opp,out}}\big), \tag{A.10}$$

which Corollary 1 signs.

Corollary 1. *In the symmetric benchmark $z_i = 0$, for any $q > 0$: $\mathcal{A}_i > 0$ for all $\beta_i > 0$, and $\mathcal{A}_i$ is strictly increasing in β_i.*

Proof. Write $h(\beta) \equiv g_q(0) - \frac{1}{2}\big[g_q(\beta) + g_q(-\beta)\big]$, so $\mathcal{A}_i = h(\beta_i)$ and $h(0) = 0$. Since ϕ is continuous, g_q is differentiable with $g_q'(m) = \phi(m) - \phi(m-q)$, and by symmetry $\phi(-x) = \phi(x)$,

$$g_q'(\beta) - g_q'(-\beta) = \big[\phi(\beta) - \phi(\beta - q)\big] - \big[\phi(\beta) - \phi(\beta + q)\big] = \phi(\beta + q) - \phi(\beta - q),$$

so that

$$h'(\beta) \;=\; -\tfrac{1}{2}\big[g_q'(\beta) - g_q'(-\beta)\big] \;=\; \tfrac{1}{2}\big[\phi(\beta - q) - \phi(\beta + q)\big].$$

For $\beta > 0$ and $q > 0$, $|\beta - q| < \beta + q$: if $\beta \geq q$ then $\beta - q < \beta + q$, and if $\beta < q$ then $q - \beta < q + \beta$. Since ϕ is strictly decreasing in $|\cdot|$ (Assumption A.1), $\phi(\beta - q) > \phi(\beta + q)$, so $h'(\beta) > 0$ for all $\beta > 0$. With $h(0) = 0$ this gives $h(\beta) > 0$ and h strictly increasing on

$(0, \infty)$. □

Corollary 2. *Conditional on z_i and q, the same-party penalty is identical whether the two candidates belong to voter i's own party or to the opposing party.*

Proof. In either case the common party term cancels and $m_i^{\text{same}} = z_i$, so the penalty is $g_q(z_i)$. □

A.6 Extensions and remarks

Remark A.1 (Amplification away from the symmetric benchmark)**.** Corollary 1 is stated at $z_i = 0$, but the conclusion does not depend on that normalization. Fix $q > 0$ and $\beta_i > 0$ and read amplification as a function of the non-party margin,

$$\mathcal{A}_i(z) \;\equiv\; g_q(z) - \tfrac{1}{2}\big[g_q(z+\beta_i) + g_q(z-\beta_i)\big].$$

Then $\mathcal{A}_i(z) > 0$ for every $z \in [0, q]$, and $\mathcal{A}_i(q - z) = \mathcal{A}_i(z)$.

Proof. Let $D(x, s) \equiv \phi(x + s) - \phi(x - s)$. Since ϕ is strictly decreasing in $|\cdot|$, for $s > 0$ we have $D(x, s) > 0$ if and only if $|x + s| < |x - s|$, that is, if and only if $x < 0$. Writing $\psi(s) \equiv \frac{1}{2}[g_q(z + s) + g_q(z - s)]$ we have $\mathcal{A}_i(z) = -\int_0^{\beta_i} \psi'(s)\,\mathrm{d}s$, and substituting $g_q'(m) = \phi(m) - \phi(m - q)$ and collecting terms gives

$$\mathcal{A}_i(z) \;=\; \tfrac{1}{2}\int_0^{\beta_i} \big[D(z - q,\, s) - D(z,\, s)\big]\,\mathrm{d}s.$$

If $z \in [0, q]$ then $z - q \le 0 \le z$, so the first term is nonnegative and the second nonpositive, with strict inequality on a set of positive measure; hence $\mathcal{A}_i(z) > 0$. Symmetry about $q/2$ follows from $g_q(m) = G(|m - q/2|)$. □

Setting $z = 0$ recovers Corollary 1. The interval $[0, q]$ is exactly the range of non-party margins over which the scandal is capable of reversing the voter's ranking: below 0 the voter rejects the candidate even when she is clean, above q she is preferred even when scandalized,

and in both cases the scandal changes nothing in any contest, so such voters contribute no punishment to any of the three penalties. The bound is conservative rather than tight—by continuity $\mathcal{A}_i > 0$ on a strictly larger neighborhood, whose extent depends on ϕ.

Remark A.2 (Amplification is non-monotone in severity)**.** Fix $z_i = 0$ and $\beta_i > 0$, and read $\mathcal{A}_i$ as a function of the scandal cost q. Then

$$\mathcal{A}_i(0) = 0, \qquad \mathcal{A}_i(q) > 0 \text{ for all } q > 0, \qquad \lim_{q\to\infty} \mathcal{A}_i(q) = 0,$$

so $\mathcal{A}_i$ attains an interior maximum in q and is neither increasing nor decreasing throughout. The first two claims are immediate from the definition and from Corollary 1. For the third, $g_q(0) \to \Phi(0) = \frac{1}{2}$, $g_q(\beta_i) \to \Phi(\beta_i)$, and $g_q(-\beta_i) \to \Phi(-\beta_i)$ as $q \to \infty$, and $\Phi(\beta_i)+\Phi(-\beta_i) = 1$ by symmetry, so $\mathcal{A}_i \to \frac{1}{2} - \frac{1}{2} = 0$. Substantively: a disqualifying scandal costs the candidate every voter who preferred her, in either type of contest, so composition ceases to matter. The local comparative static is

$$\frac{\partial \mathcal{A}_i}{\partial q} = \phi(q) - \tfrac{1}{2}\big[\phi(q-\beta_i) + \phi(q+\beta_i)\big],$$

positive if and only if ϕ at q exceeds the average of its values at $q \pm \beta_i$. This holds for q near zero, since $\phi(0) > \phi(\beta_i)$, and fails for q large. Where it turns depends on ϕ: for the standard normal with $\beta_i = 1$ the sign changes near $q = 1.09$, while for the Laplace density it changes near $q = 0.75$.

Remark A.3 (Heterogeneous party attachment and Independents)**.** Everything in Proposition 1 and Corollary 1 operates through β_i, and $\mathcal{A}_i$ is strictly increasing in it. Aggregate amplification is therefore increasing in the share of strongly attached partisans, and vanishes for pure Independents: $\beta_i = 0$ gives $m^{\text{same}} = m^{\text{opp}} = z_i$, so composition is irrelevant to them. Two implications follow. First, excluding Independents should widen the estimated inter/intra separation, which it does (Section 6.1). Second, as β_i grows, $\frac{1}{2}[g_q(\beta_i)+g_q(-\beta_i)] \to 0$

while $g_q(z_i)$ is unchanged: polarization erodes inter-party accountability and leaves intra-party accountability intact.

Remark A.4 (Structural mapping)**.** Equation (9) in the main text is the empirical counterpart of (A.3) under three restrictions: common coefficients ($\beta_i = \beta$, $\gamma_i = \gamma$), the ideological term folded into the match coefficients (the estimated β is total alignment utility, not an identity premium net of spatial distance), and, in the pooled version, an additive binary valence function. The type-specific models in Section 7 relax the last restriction.

B Experimental Design

B.1 Conjoint attributes and wording

Table B1: Conjoint attributes and structural levels.

Attribute	Levels
Gender	Male; Female
Party affiliation	Democratic; Republican
Faction	Democratic candidates: Progressive; Moderate. Republican candidates: MAGA; Moderate.
Political scandal	No scandal; corruption; credentials; sexual misconduct; violent protest.
Positive valence	No positive valence; workplace respect; fact-checking accuracy; prevented layoffs; community service.

Note: Profiles were presented as short narrative descriptions. The order of the scandal and positive-valence sentences was randomized. Omitted levels correspond to "No scandal" and "No positive valence."

Table B2: Exact wording of scandal and positive-valence levels.

Category	Exact wording shown to respondents
No scandal	No scandal.
Corruption scandal	"Media reports revealed that the candidate had accepted $500,000 in illegal campaign donations."
Credentials scandal	"Media reports revealed that the candidate fabricated having a university degree and misrepresented their educational background on their CV."
Sexual misconduct scandal	"Media reports revealed that the candidate had sent unwanted sexually explicit text messages to a staff member."
Violent protest scandal	"Media reports revealed that the candidate had taken part in a violent riot involving acts of vandalism and physical confrontations while underage."
No positive valence	No positive valence.
Workplace respect cue	"The candidate was widely praised for enforcing strict anti-harassment policies and fostering a respectful workplace."
Fact-checking cue	"The candidate had 95% of campaign statements certified as accurate by an independent fact checker."
Prevented layoffs cue	"The candidate led a public-private partnership that prevented layoffs during a local economic downturn."
Community service cue	"The candidate received a national award for community service while underage."

B.2 Outcome and treatment coding

Table B3: Outcome wording and coding.

Outcome	Survey task wording	Coding
Vote choice	Respondents chose which of the two candidates they would vote for in the paired task.	Profile-level indicator coded 1 if the candidate was chosen and 0 otherwise.
Competence	After choosing, respondents rated each candidate's competence.	Candidate-level rating on a 0–10 scale.
Honesty	After choosing, respondents rated each candidate's honesty.	Candidate-level rating on a 0–10 scale.

Table B4: Contest-type coding.

Category	Definition
Inter-party contest	The two profiles in a task have different party labels.
Intra-party contest	The two profiles in a task share the same party label.
Out-group intra-party contest	An intra-party contest in which both candidates belong to the party opposite the respondent's party camp.
In-group intra-party contest	An intra-party contest in which both candidates belong to the respondent's own party camp.

Table B5: Faction-alignment coding.

Category	Definition
Strict co-faction / rival-faction	Restricted to Democratic and Republican identifiers, co-partisan candidate profiles, and intra-party contests. "Co-faction" means the respondent and candidate share the same within-party faction; "Rival faction" means they do not.
Inclusive four-camp coding	All respondents are retained. Respondents are assigned to one of four camps using party identification and the faction-closeness item: Moderate Democrat, Progressive Democrat, Non-MAGA Republican, and MAGA Republican. Candidate profiles are mapped into the same four camps, with Republican Moderate profiles corresponding to the respondent-side Non-MAGA Republican camp.
Inclusive alignment categories	Same faction camp; same party camp, rival faction; opposing party camp.

B.3 Sample and descriptive statistics

Table B6: Survey quota targets and achieved sample shares.

Category	Target	Achieved
Total sample	3,000	3,000
Male	49%	48%
Female	51%	51%
Other	–	1%
Age 18–24	13%	13%
Age 25–34	19%	19%
Age 35–44	19%	19%
Age 45–54	17%	17%
Age 55–75	32%	32%
Northeast	17%	17%
Midwest	20%	20%
South	39%	39%
West	23%	23%

Note: Quota targets and achieved shares are reported from the survey provider's quota summary for the conjoint sample. A dash indicates that no target quota was set for that category.

Table B7: Descriptive counts by contest type.

Contest type	Tasks	Profile observations
Inter-party contest	6089	12178
Intra-party contest	5911	11822
Out-group intra-party contest	2964	5928
In-group intra-party contest	2947	5894

Table B8: Respondent party identification shares.

Respondent party	Respondents	Share
Democratic	1,167	0.389
Republican	970	0.323
Independent	863	0.288

Table B9: Respondent thermometer means and medians.

Item	Mean	Median
Democratic Party thermometer	46.847	50.000
Republican Party thermometer	40.308	34.000
Trump thermometer	36.375	14.000
AOC thermometer	40.500	45.000

C Main Model Details

The main profile-level specification is estimated with respondent-clustered standard errors. Table C1 reports a compact version of the main vote model. The remaining tables show that the estimated penalties are stable without controls, in a logit specification reported as average marginal effects by contest type, and when the partisan-strength gap control is modeled quadratically or in bins. We do not report respondent-task fixed effects as a separate table because those estimates are not directly comparable to the profile-level average-comparison estimates in the main text; the paired-loss estimand provides a more transparent within-task check that treats each conjoint pair as the relevant choice set.

C.1 Main specification

Table C1: Compact H1 vote model.

Term	Estimate	Std. error	95% CI
Scandal	-0.187	0.012	[-0.210, -0.164]
Intra-party contest	0.075	0.012	[0.050, 0.099]
Scandal x intra-party	-0.106	0.015	[-0.136, -0.076]
Profile controls		Yes	
Respondent controls		Yes	
Task fixed effects		Yes	
Standard errors clustered by respondent		Yes	
Profile observations		24,000	
Respondents		3,000	

C.2 Robustness to model specification

Table C2: Vote penalties by contest type across robustness specifications.

Model	Contest type	Penalty	95% CI low	95% CI high
Main LPM	Inter-party contest	-0.187	-0.210	-0.164
Main LPM	Intra-party contest	-0.292	-0.313	-0.272
No-controls LPM	Inter-party contest	-0.194	-0.217	-0.171
No-controls LPM	Intra-party contest	-0.284	-0.305	-0.264
Respondent FE LPM	Inter-party contest	-0.211	-0.237	-0.185
Respondent FE LPM	Intra-party contest	-0.331	-0.354	-0.308
Logit AME	Inter-party contest	-0.187	-0.210	-0.163
Logit AME	Intra-party contest	-0.292	-0.312	-0.271
Quadratic strength	Inter-party contest	-0.187	-0.210	-0.164
Quadratic strength	Intra-party contest	-0.293	-0.313	-0.272
Binned strength	Inter-party contest	-0.187	-0.210	-0.164
Binned strength	Intra-party contest	-0.293	-0.313	-0.272

Table C3: Vote intra-party amplification contrast across robustness specifications.

Specification	Estimate	95% CI
Main LPM	-0.106	[-0.136, -0.076]
No-controls LPM	-0.090	[-0.121, -0.060]
Respondent FE LPM	-0.120	[-0.154, -0.087]
Quadratic strength	-0.106	[-0.136, -0.076]
Binned strength	-0.106	[-0.136, -0.076]

C.3 Explicit contrast tables

Tables C4 and C5 report the direct contrasts behind the main intra-party amplification result and the respondent-relation scope test.

Table C4: Main scandal penalty and intra-party amplification contrast.

Contrast	Estimate	CI low	CI high
Inter-party scandal penalty	-0.187	-0.210	-0.164
Intra-party scandal penalty	-0.292	-0.313	-0.272
Intra-party minus inter-party	-0.106	-0.136	-0.076

Note: Negative estimates indicate lower vote support. The final row is the directly estimated interaction contrast.

Table C5: Respondent-relation scandal penalties and contrasts.

Contrast	Estimate	CI low	CI high
Inter-party scandal penalty	-0.187	-0.210	-0.164
Out-group intra-party scandal penalty	-0.275	-0.304	-0.247
In-group intra-party scandal penalty	-0.310	-0.338	-0.281
Out-group intra-party minus inter-party	-0.089	-0.125	-0.053
In-group intra-party minus inter-party	-0.123	-0.159	-0.087
In-group minus out-group intra-party	-0.034	-0.074	0.006

Note: The last three rows are directly estimated contrasts among the context-specific scandal penalties.

C.4 Secondary evaluation outcomes

Vote choice is the primary outcome used to test the paper's hypotheses. The experiment also collected 0–10 competence and honesty ratings after respondents made each vote choice; Table B3 reports their wording and coding. We use these ratings only to assess whether the electoral patterns are accompanied by changes in candidate evaluations.

Table C6 shows that the pooled intra-party amplification also appears in both ratings. Table C7 then separates the three contest contexts. The competence and honesty penalties change little between inter-party and out-group intra-party contests, and both out-group-minus-inter-party intervals include zero. The additional evaluative penalty is instead concentrated in in-group intra-party contests. This divergence is compatible with a choice-set

mechanism in which changing the available alternative can change vote choice without requiring an equally large change in the perceived attributes of the scandalized candidate. Because the ratings were elicited after the vote, however, they are supplementary outcomes rather than independent measures of the process producing that vote.

Table C6: Scandal penalties and intra-party amplification across outcomes.

Outcome	Inter-party penalty	Intra-party penalty	Intra-party minus inter-party amplification contrast
Vote choice	−0.187 [−0.210, −0.164]	−0.292 [−0.313, −0.272]	−0.106 [−0.136, −0.076]
Competence	−1.472 [−1.613, −1.331]	−1.780 [−1.909, −1.650]	−0.308 [−0.486, −0.129]
Honesty	−1.924 [−2.066, −1.782]	−2.327 [−2.464, −2.190]	−0.403 [−0.583, −0.223]

Note: Entries are average comparisons for a change from no scandal to any scandal; brackets contain 95% confidence intervals. Vote effects are changes in the predicted probability that a profile is chosen. Competence and honesty effects are changes on the corresponding 0–10 rating scales. Each outcome is estimated in a separate profile-level model with the same candidate attributes, respondent covariates, and task-order fixed effects as Equation (6); standard errors are clustered by respondent.

Table C7: Scandal penalties across contest contexts and outcomes.

Panel A: Estimated scandal penalties

Outcome	Inter-party	Out-group intra-party	In-group intra-party
Vote choice	−0.187 [−0.210, −0.164]	−0.275 [−0.304, −0.247]	−0.310 [−0.338, −0.281]
Competence	−1.478 [−1.619, −1.337]	−1.549 [−1.729, −1.370]	−2.036 [−2.199, −1.872]
Honesty	−1.930 [−2.072, −1.788]	−2.050 [−2.234, −1.866]	−2.628 [−2.811, −2.445]

Panel B: Differences between contest contexts

Outcome	Out-group minus inter-party	In-group minus inter-party	In-group minus out-group
Vote choice	−0.089 [−0.125, −0.053]	−0.123 [−0.159, −0.087]	−0.034 [−0.074, 0.006]
Competence	−0.071 [−0.289, 0.147]	−0.558 [−0.763, −0.352]	−0.486 [−0.722, −0.251]
Honesty	−0.120 [−0.338, 0.098]	−0.699 [−0.917, −0.480]	−0.579 [−0.831, −0.327]

Note: Brackets contain 95% confidence intervals. Panel A reports the average comparison for scandal within each contest context. Panel B reports differences between those penalties. Vote effects are expressed as changes in choice probability; competence and honesty effects are expressed in 0–10 rating units. Each outcome is estimated in a separate profile-level model with respondent-clustered standard errors. Independents are assigned to a party camp using the pre-treatment faction-closeness item for the out-group/in-group classification.

D Scandal-Type and Valence Analyses

D.1 Scandal-type penalties

Table D1: Vote-choice scandal penalties by type and contest composition.

Scandal type	Inter-party	Intra-party
Credentials	-0.164	-0.254
Corruption	-0.199	-0.330
Sexual misconduct	-0.254	-0.379
Violent protest	-0.127	-0.220

Note: Entries are profile-level average comparisons relative to "No scandal." Figure 6 shows the corresponding confidence intervals.

Table D2: Scandal-type vote penalties and amplification contrasts.

Contrast	Estimate	CI low	CI high
Credentials: inter-party penalty	-0.164	-0.191	-0.136
Credentials: intra-party penalty	-0.254	-0.281	-0.227
Credentials: intra-party minus inter-party	-0.090	-0.128	-0.052
Corruption: inter-party penalty	-0.199	-0.230	-0.169
Corruption: intra-party penalty	-0.330	-0.357	-0.303
Corruption: intra-party minus inter-party	-0.131	-0.171	-0.091
Sexual misconduct: inter-party penalty	-0.254	-0.282	-0.225
Sexual misconduct: intra-party penalty	-0.379	-0.407	-0.350
Sexual misconduct: intra-party minus inter-party	-0.125	-0.164	-0.087
Violent protest: inter-party penalty	-0.127	-0.156	-0.097
Violent protest: intra-party penalty	-0.220	-0.249	-0.192
Violent protest: intra-party minus inter-party	-0.094	-0.134	-0.054

Note: For each scandal type, the amplification row is the directly estimated intra-party minus inter-party contrast.

D.2 Paired-loss estimates by scandal type

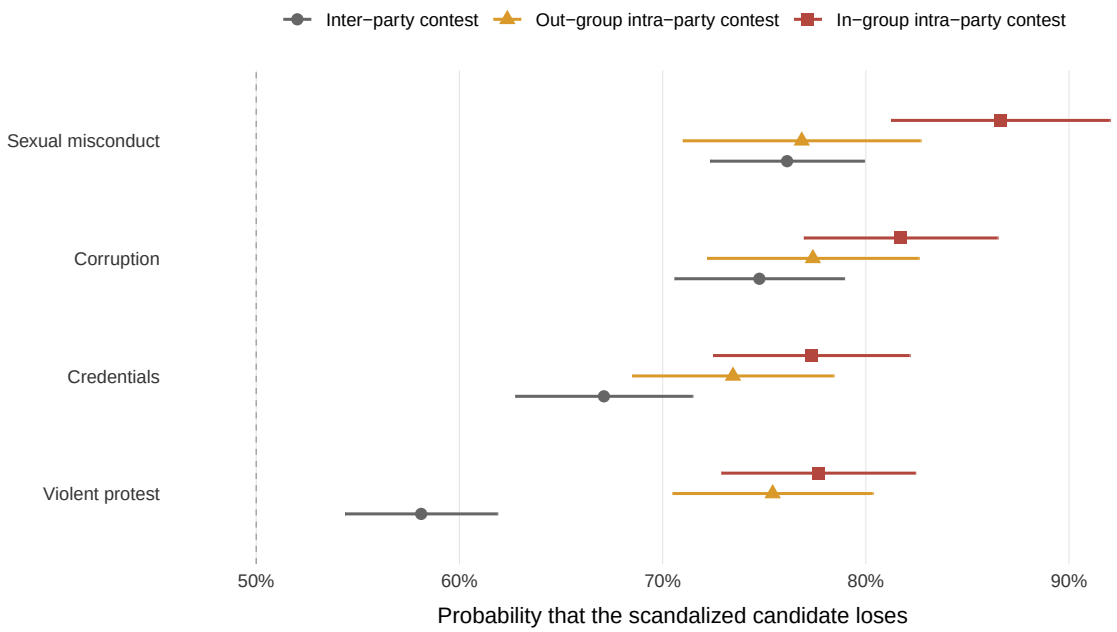


Figure D1: Paired vote-loss probabilities by scandal type and respondent relation to the candidate pair.

Note: The sample is restricted to tasks with exactly one scandalized profile. Points report the probability that the scandalized candidate is not chosen, separately by scandal type and contest context; horizontal bars are 95% confidence intervals with standard errors clustered by respondent.

Table D3: Paired vote loss probabilities by scandal type and respondent relation to the candidate pair.

Scandal type	Inter-party contest	Out-group intra-party	In-group intra-party
Violent protest	0.581	0.754	0.777
Credentials	0.671	0.735	0.773
Corruption	0.748	0.774	0.817
Sexual misconduct	0.761	0.768	0.866

Note: Entries are point estimates. Figure D1 has confidence intervals.

D.3 Two-scandal severity ranking

Table D4: Relative scandal severity in intra-party two-scandal vote tasks.

Scandal type	Avg. loss prob.	95% CI	Profile obs.	Respondents
Sexual misconduct	0.607	[0.588, 0.626]	3,061	2,101
Corruption	0.531	[0.510, 0.552]	2,429	1,799
Credentials	0.455	[0.436, 0.475]	2,701	1,830
Violent protest	0.395	[0.376, 0.414]	2,671	1,951

Note: Entries report the probability that a given scandal loses when paired against another scandalized candidate in intra-party vote tasks.

D.4 Positive-valence cues

The experiment also randomized positive accomplishments. Positive-valence cues improve candidate support, and their pooled effect is larger in intra-party than in inter-party contests, although cue-specific amplification contrasts are estimated less precisely. This parallel pattern is consistent with the model's general pivotality logic: removing party-based differentiation makes both positive and negative valence more consequential. The smaller absolute amplification of positive cues is broadly consistent with their smaller estimated utility weights. Figure D2 and Table D5 report cue-specific vote effects by contest type, while Table D6 reports the pooled and cue-specific amplification contrasts. Table D7 reports an exploratory joint status analysis distinguishing neutral, positive-only, negative-only, and mixed profiles.

Table D5: Positive-valence vote effects by contest type.

Positive-valence cue	Inter-party	Intra-party
Workplace respect	0.106	0.142
Fact-checking accuracy	0.072	0.113
Prevented layoffs	0.092	0.134
Community service	0.049	0.063

Note: Entries are profile-level average comparisons relative to "No positive valence." Figure D2 shows the corresponding confidence intervals.

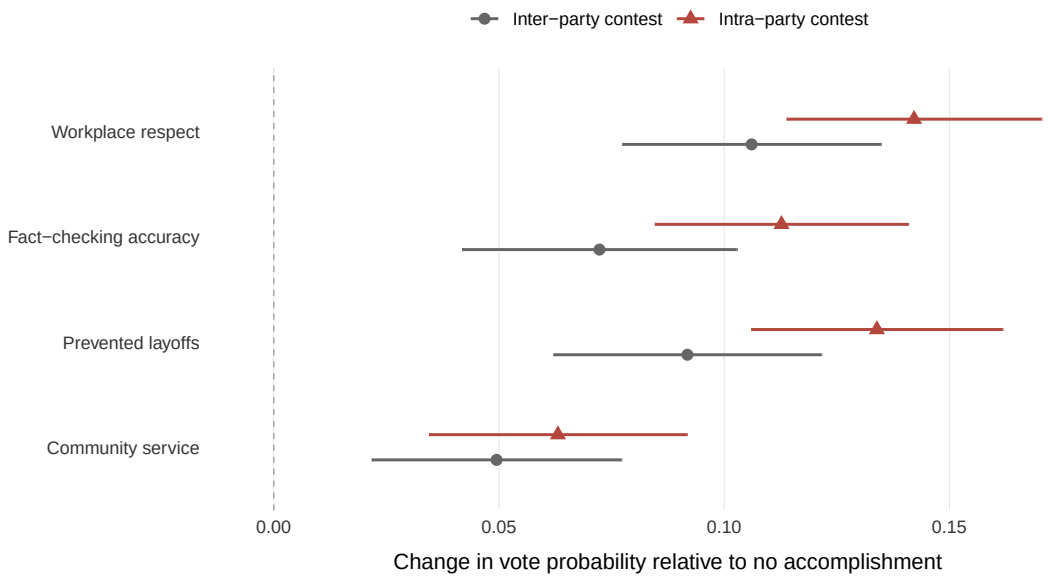


Figure D2: Positive-valence vote effects by contest type.

Note: Points report profile-level average comparisons: the change in predicted vote probability from no positive information to the indicated accomplishment, estimated within contest type and averaged over the remaining variables. Horizontal bars are 95% confidence intervals with standard errors clustered by respondent.

Table D6: Positive-valence vote effects and amplification contrasts.

Contrast	Estimate	CI low	CI high
Pooled any-positive-cue contrast			
Pooled positive cue: inter-party effect	0.079	0.056	0.103
Pooled positive cue: intra-party effect	0.115	0.093	0.138
Pooled positive cue: intra-party minus inter-party	0.036	0.004	0.068
Cue-specific contrasts			
Workplace respect: inter-party effect	0.106	0.077	0.135
Workplace respect: intra-party effect	0.142	0.114	0.170
Workplace respect: intra-party minus inter-party	0.036	-0.004	0.076
Fact-checking accuracy: inter-party effect	0.072	0.042	0.103
Fact-checking accuracy: intra-party effect	0.113	0.085	0.141
Fact-checking accuracy: intra-party minus inter-party	0.040	-0.001	0.082
Prevented layoffs: inter-party effect	0.092	0.062	0.122
Prevented layoffs: intra-party effect	0.134	0.106	0.162
Prevented layoffs: intra-party minus inter-party	0.042	0.002	0.082
Community service: inter-party effect	0.049	0.022	0.077
Community service: intra-party effect	0.063	0.034	0.092
Community service: intra-party minus inter-party	0.014	-0.026	0.054

Note: Positive estimates indicate higher vote support relative to no positive cue. The pooled rows use a binary indicator for any of the four positive cues. Amplification rows are direct intra-party-minus-inter-party contrasts with respondent-clustered confidence intervals.

Table D7: Exploratory joint valence-status effects on vote choice.

Status	Estimand	Estimate	CI low	CI high
Panel A: Effects relative to neutral profiles				
Positive only	inter-party effect	0.097	0.044	0.149
Positive only	intra-party effect	0.112	0.064	0.161
Positive only	intra-party minus inter-party	0.016	-0.054	0.086
Negative only	inter-party effect	-0.159	-0.212	-0.107
Negative only	intra-party effect	-0.290	-0.340	-0.240
Negative only	intra-party minus inter-party	-0.130	-0.202	-0.059
Mixed	inter-party effect	-0.092	-0.143	-0.042
Mixed	intra-party effect	-0.175	-0.220	-0.131
Mixed	intra-party minus inter-party	-0.083	-0.150	-0.016

Panel B: Direct compensation, scandal, and additivity contrasts

Contrast	Contest	Estimate	SE	95% CI	p
Mixed − Negative only	Inter-party	0.067	0.013	[0.042, 0.092]	< 0.001
Mixed − Negative only	Intra-party	0.114	0.013	[0.088, 0.140]	< 0.001
Mixed − Positive only	Inter-party	−0.189	0.013	[−0.213, −0.164]	< 0.001
Mixed − Positive only	Intra-party	−0.288	0.012	[−0.311, −0.264]	< 0.001
Positive only − Neutral	Inter-party	0.097	0.027	[0.044, 0.149]	< 0.001
Positive only − Neutral	Intra-party	0.112	0.025	[0.064, 0.161]	< 0.001
Negative only − Neutral	Inter-party	−0.159	0.027	[−0.212, −0.107]	< 0.001
Negative only − Neutral	Intra-party	−0.290	0.026	[−0.340, −0.240]	< 0.001
Additivity residual	Inter-party	−0.029	0.028	[−0.085, 0.026]	0.303
Additivity residual	Intra-party	0.002	0.029	[−0.055, 0.058]	0.952
Intra-party − inter-party residual	Difference	0.031	0.040	[−0.047, 0.109]	0.435

Note: Panel A reports effects relative to neutral profiles with no scandal and no positive cue. Panel B reports covariance-based linear contrasts from the same joint model, with standard errors clustered by respondent. Mixed − Negative only is the positive-cue effect among scandalized profiles; Mixed − Positive only is the scandal effect when a positive cue is present. The additivity residual is (Mixed − Negative only) − (Positive only − Neutral), algebraically identical to (Mixed − Positive only) − (Negative only − Neutral). The final row compares this residual across contest types.

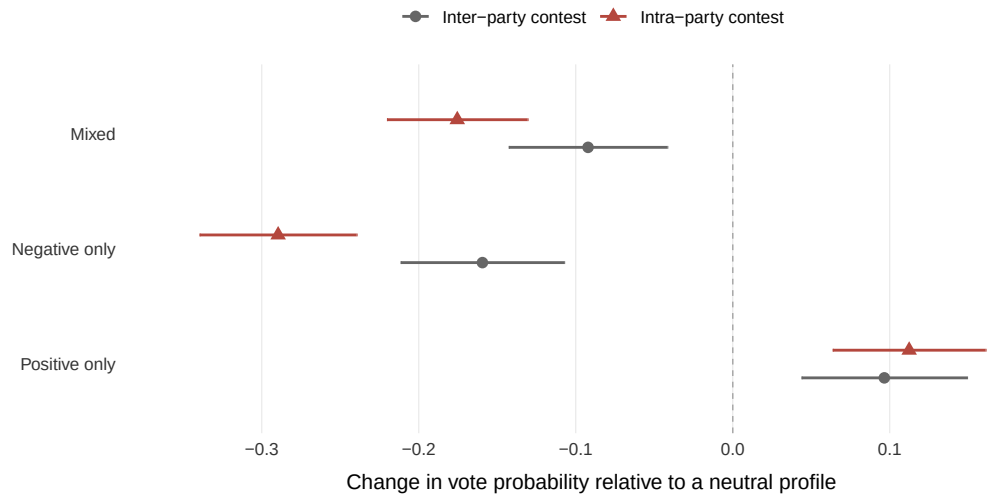


Figure D3: Exploratory joint valence-status effects by contest type.

Note: Reference category is a neutral profile with no scandal and no positive cue. The model includes valence-status category, contest type, their interaction, candidate controls, respondent controls, task-order fixed effects, and respondent-clustered standard errors.

E Party and Faction Scope Conditions

E.1 Respondent relation to the candidate pair

Table C7 reports the corresponding three-context estimates for vote choice and the secondary ratings. The four-context split among Democratic and Republican identifiers is reported in the main text (Table 2). Its inter-party side split is less theoretically clean than the pooled inter-party baseline, and we read it as a mechanism check rather than a stand-alone result.

E.2 Party-family decomposition

Table E1: Vote-choice scandal penalties by party family.

Context	Penalty	95% CI low	95% CI high
Inter-party contest	-0.184	-0.207	-0.161
Democratic intra-party contest	-0.305	-0.333	-0.277
Republican intra-party contest	-0.296	-0.326	-0.267

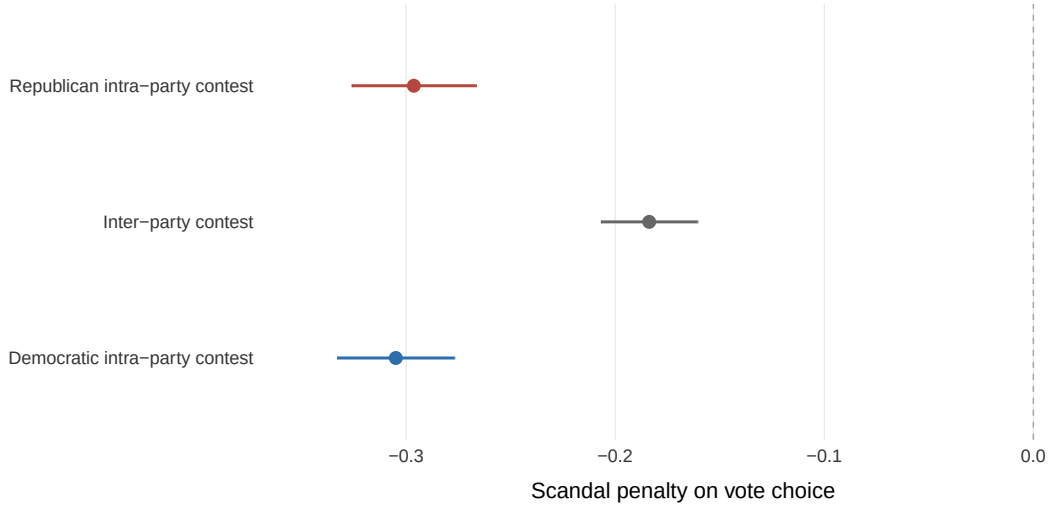


Figure E1: Vote penalties by party family of the intra-party contest.

Note: Points report profile-level average comparisons for scandal in inter-party contests and in same-party contests composed of Democratic or Republican candidates. Horizontal bars are 95% confidence intervals with standard errors clustered by respondent.

Table E2: Party-family scandal penalties and contrasts.

Contrast	Estimate	CI low	CI high
Inter-party scandal penalty	-0.184	-0.207	-0.161
Democratic intra-party scandal penalty	-0.305	-0.333	-0.277
Republican intra-party scandal penalty	-0.296	-0.326	-0.267
Democratic intra-party minus inter-party	-0.121	-0.156	-0.086
Republican intra-party minus inter-party	-0.113	-0.150	-0.075
Democratic minus Republican intra-party	-0.009	-0.049	0.032

Note: Democratic and Republican intra-party rows classify the party family of the same-party candidate pair.

E.3 Strict factional alignment

The strict co-faction versus rival-faction comparison is secondary to the party-composition tests. The figure and corresponding estimates are reported below.

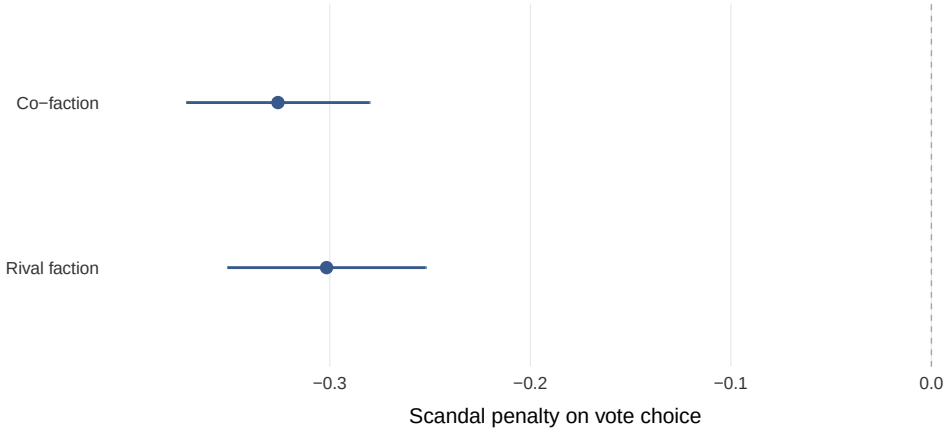


Figure E2: Strict co-faction versus rival-faction vote penalties in intra-party contests.

Note: Estimand: scandal penalty on vote within co-partisan intra-party contests, by whether the respondent shares the candidate's faction. Restricted to Democratic and Republican identifiers. Error bars: 95% confidence intervals. SEs clustered by respondent.

Table E3: Strict co-faction versus rival-faction scandal penalties.

Faction alignment	Penalty	95% CI low	95% CI high
Co-faction	-0.326	-0.371	-0.280
Rival faction	-0.302	-0.351	-0.252

E.4 Factional configuration of the candidate pair

The strict alignment analysis above classifies the focal candidate relative to the respondent. Because faction labels were randomized independently across profiles, we can also distinguish the factional alignment of the available alternative. Panel A of Table E4 reports the probability that the scandalized candidate loses in own-party intra-party tasks with exactly one scandal; Panel B reports the incremental profile-level scandal penalty using all own-party intra-party tasks. In both panels, A1 changes the alternative's alignment while holding the focal or scandalized candidate co-factional, A2 makes the same change while holding that

candidate rival-factional, and A1−A2 directly tests whether the alternative-faction contrast differs across those two configurations.

Table E4: Factional configuration of the scandalized or focal candidate and the available alternative.

Configuration or contrast	Estimate	SE	95% CI	p	Respondents	Tasks
Panel A: Probability that the scandalized candidate loses						
Scandalized co-faction; clean alternative co-faction	0.824	0.026	[0.772, 0.875]	–	197	204
Scandalized co-faction; clean alternative rival-faction	0.709	0.039	[0.633, 0.784]	–	143	151
Scandalized rival-faction; clean alternative co-faction	0.816	0.028	[0.761, 0.870]	–	193	206
Scandalized rival-faction; clean alternative rival-faction	0.827	0.027	[0.775, 0.880]	–	194	197
A1: clean co-faction minus rival-faction, scandalized co-faction	0.115	0.046	[0.024, 0.205]	.013	–	–
A2: clean co-faction minus rival-faction, scandalized rival-faction	−0.012	0.038	[−0.087, 0.063]	.755	–	–
A1−A2	0.127	0.059	[0.011, 0.242]	.032	–	–
Panel B: Incremental profile-level scandal penalty						
Configuration or contrast	Estimate	SE	95% CI	p	Respondents	Profiles
Focal co-faction; alternative co-faction	−0.345	0.030	[−0.403, −0.287]	–	444	958
Focal co-faction; alternative rival-faction	−0.267	0.031	[−0.327, −0.206]	–	887	1,135
Focal rival-faction; alternative co-faction	−0.278	0.036	[−0.348, −0.208]	–	887	1,135
Focal rival-faction; alternative rival-faction	−0.341	0.028	[−0.396, −0.286]	–	498	1,072
A1: alternative co-faction minus rival-faction, focal co-faction	−0.078	0.043	[−0.163, 0.007]	.073	–	–
A2: alternative co-faction minus rival-faction, focal rival-faction	0.063	0.045	[−0.025, 0.152]	.162	–	–
A1−A2	−0.141	0.061	[−0.261, −0.021]	.021	–	–

Note: Democratic and Republican identifiers evaluating two candidates from their own party. Panel A is restricted to tasks with exactly one scandalized profile; its estimate is the probability that the scandalized candidate loses, and counts refer to respondents and tasks. Panel B uses all own-party intra-party tasks; its estimate is the change in the focal profile's choice probability when it changes from clean to scandalized, and the final count column reports profiles rather than tasks. Negative profile estimates indicate larger scandal penalties. Both models use the jointly randomized faction labels of the two profiles. Standard errors and direct contrasts use respondent-clustered covariance matrices; 95% confidence intervals use the normal approximation.

The direct interaction is 12.7 percentage points for paired loss and 14.1 points on a punishment-magnitude scale for the profile-level penalty. The two profile-level component contrasts are individually imprecise, however, and their directions differ: the result is a configuration interaction rather than evidence that a co-factional alternative uniformly increases

punishment.

Table E5: Observed and structural-model-implied scandal penalties by factional configuration.

Configuration or contrast	Observed	95% CI	Model-implied	95% CI	Obs. − model
Focal co-faction; alternative co-faction	−0.345	[−0.403, −0.287]	−0.260	[−0.278, −0.243]	−0.084
Focal co-faction; alternative rival-faction	−0.267	[−0.327, −0.206]	−0.277	[−0.298, −0.257]	0.010
Focal rival-faction; alternative co-faction	−0.278	[−0.348, −0.208]	−0.233	[−0.250, −0.216]	−0.045
Focal rival-faction; alternative rival-faction	−0.341	[−0.396, −0.286]	−0.261	[−0.278, −0.243]	−0.080
A1: alternative co-faction minus rival-faction, focal co-faction	−0.078	[−0.163, 0.007]	0.017	[0.012, 0.021]	−0.095
A2: alternative co-faction minus rival-faction, focal rival-faction	0.063	[−0.025, 0.152]	0.028	[0.018, 0.038]	0.035
A1−A2	−0.141	[−0.261, −0.021]	−0.011	[−0.018, −0.005]	−0.130

Note: Observed estimates reproduce Panel B of Table E4. Model-implied penalties come from the existing partisan-only structural calibration ($\beta = 1.449$, $\gamma = 0.343$, $\kappa = -1.172$, and $\rho = 0.445$), preserving observed profile position and positive-valence differences. The model contains no scandal-by-faction interaction; configuration affects the penalty only through faction's contribution to the clean utility margin and the nonlinear logit link. Model-implied intervals use the delta method. Observed-minus-model residuals are descriptive because both estimates use overlapping data and their joint covariance is not estimated.

Table E5 applies the paper's existing additive structural calibration to the same four configurations while preserving the observed distribution of the other included attributes. The model adds no scandal-by-faction parameter: faction alignment shifts the clean utility margin by γ, and the nonlinear choice probability translates that shift into a model-implied penalty. It predicts only a small A1−A2 interaction and does not reproduce the observed configuration pattern, so the design-based results are treated as exploratory rather than as confirmation of a general factional-proximity mechanism.

E.5 Inclusive factional alignment

Table E6: Inclusive faction-alignment scandal penalties.

Faction alignment	Penalty	95% CI low	95% CI high
Opposing party camp	-0.257	-0.279	-0.235
Same party camp, rival faction	-0.232	-0.259	-0.204
Same faction camp	-0.227	-0.252	-0.201

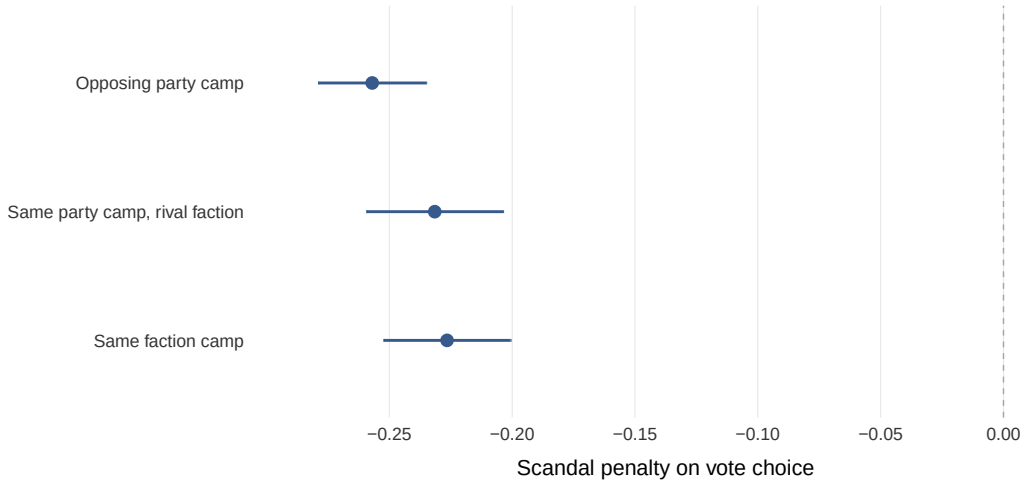


Figure E3: Inclusive faction-alignment vote penalties.

Note: Points report profile-level average comparisons for scandal under the inclusive four-camp alignment coding defined in Appendix B. Horizontal bars are 95% confidence intervals with standard errors clustered by respondent.

E.6 Cell counts and respondent-party heterogeneity

Table E7: Factional cell counts for appendix scope-condition analyses.

Sample	Respondents	Tasks	Profiles
Full conjoint sample	3000	12000	24000
Strict co-faction sample	1387	2150	4300
Co-faction	1107	1569	2003
Rival faction	1218	1716	2297
Inclusive faction sample	3000	12000	24000
Same faction camp	2595	5285	5976
Same party camp, rival faction	2602	5289	6007
Opposing party camp	3000	9053	12017
Democratic Moderate	1024	4096	8192
Democratic Progressive	603	2412	4824
Republican Moderate	707	2828	5656
Republican MAGA	666	2664	5328

Note: Respondent and task counts in the alignment rows are not mutually exclusive because a respondent or task may contribute profiles to more than one alignment category. Profile counts partition the relevant profile sample.

Table E8: Vote scandal penalties by respondent party.

Respondent party	Inter-party contest	Intra-party contest	Amplification contrast
Democratic	-0.184	-0.251	-0.068
Republican	-0.133	-0.350	-0.217
Independent	-0.243	-0.298	-0.055

Table E9: Strict co-faction versus rival-faction vote penalties by respondent party.

Respondent party	Faction alignment	Penalty	95% CI low	95% CI high
Democratic	Co-faction	-0.329	-0.389	-0.269
Democratic	Rival faction	-0.292	-0.357	-0.227
Republican	Co-faction	-0.327	-0.399	-0.256
Republican	Rival faction	-0.313	-0.392	-0.235

E.7 Partisan-strength heterogeneity

Consistent with the party-attachment mechanism, the intra-party amplification point estimates grow monotonically across partisan-strength terciles.

Table E10: Intra-party amplification by partisan-strength tercile.

Partisan-strength group	Amplification	95% CI low	95% CI high
Low partisan-strength gap	−0.059	−0.121	0.002
Medium partisan-strength gap	−0.118	−0.178	−0.058
High partisan-strength gap	−0.204	−0.267	−0.141

Note: Entries are intra-party minus inter-party scandal penalties from a profile-level vote model among Democratic and Republican identifiers. The model interacts scandal, contest type, and partisan-strength tercile, with the main candidate controls, respondent controls, task-order fixed effects, and respondent-clustered standard errors.

Table E11: Partisan-strength amplification levels and pairwise differences.

Panel	Contrast	Estimate	CI low	CI high
Amplification by tercile	Low partisan-strength gap	-0.059	-0.121	0.002
Amplification by tercile	Medium partisan-strength gap	-0.118	-0.178	-0.058
Amplification by tercile	High partisan-strength gap	-0.204	-0.267	-0.141
Pairwise differences	Medium minus low amplification	-0.059	-0.145	0.027
Pairwise differences	High minus low amplification	-0.145	-0.233	-0.057
Pairwise differences	High minus medium amplification	-0.086	-0.172	0.001

Note: The pairwise rows test whether the amplification estimates differ across terciles.

Figure 5 in the main text reports the coefficient plot; Table E10 gives the numerical

values and Table E11 the pairwise differences across terciles.

F Structural Estimation Details

F.1 Alternative treatments of Independent respondents

The main text uses strict match coding and reports its headline coefficients among Democratic and Republican identifiers. Table F1 compares the *strict*, *inclusive*, and *hybrid* codings in the full sample, so its strict column is not an exact duplicate of Table 3. Under strict coding, party and faction match are zero for Independent respondents, matching the theoretical model. The inclusive coding assigns Independents to party and faction camps using the closest-faction item; its coefficients measure camp affinity rather than co-partisanship. The hybrid coding keeps Independent party match at zero but estimates a separate closest-faction affinity for Independents, who are asked in the pre-treatment questionnaire which faction they feel closest to.

Table F1: Structural estimates under alternative treatments of Independent respondents.

Quantity	Inclusive	Strict	Hybrid
Party/camp utility	1.258	1.448	1.450
Additional faction utility	0.289	0.342	0.342
Independent closest-faction affinity	—	—	0.713
Scandal utility	−1.202	−1.166	−1.175
Positive-valence cue utility	0.447	0.441	0.449
MRS: scandal / party	0.955	0.805	0.811
MRS: scandal / faction increment	4.163	3.407	3.431
MRS: scandal / total alignment	0.777	0.651	0.656
MRS: scandal / positive-valence cue	2.687	2.641	2.617
MRS: scandal / Independent faction affinity	—	—	1.648
AIC	13,349	13,565	13,430

Note: All models use 12,000 tasks from 3,000 respondents; standard errors clustered by respondent. *Inclusive* assigns Independents to party and faction camps using the closest-faction item, so coefficients measure camp affinity rather than co-partisanship. *Strict* sets both match indicators to zero for Independents, matching the theoretical model. *Hybrid* keeps Independent party match at zero and estimates a separate closest-faction affinity. The inclusive model's lower AIC reflects the fact that it assigns Independents a non-zero regressor; it estimates a broader concept and is not comparable on fit alone.

The strict and hybrid partisan coefficients are effectively identical, and the headline MRS quantities are stable across codings. The hybrid model, which sets Independent party match to zero, estimates an Independent closest-faction affinity of 0.713 (SE 0.066), roughly half the partisans' total alignment utility. The fact that this affinity is smaller than the average scandal cost is consistent with the model's implication that voters with weaker party attachment should be more willing to sanction misconduct across party lines. This is consistent with the model's account of Independents as the voters for whom inter-party and intra-party contests least differ (Remark A.3).

F.2 Model-implied penalties by scandal type

Table F2: Model-implied penalties and increases in absolute penalty by scandal type.

Scandal	Inter-party	Out-group intra	In-group intra	Out-group increase	In-group increase
Sexual misconduct	−0.225	−0.332	−0.329	0.107	0.104
Illegal funding/corruption	−0.193	−0.298	−0.295	0.105	0.102
Falsified credentials	−0.150	−0.245	−0.243	0.095	0.093
Violent protest while underage	−0.120	−0.204	−0.202	0.084	0.082

Note: Partisan respondents, additive type model with type-specific scandal costs but no scandal-type-by-contest interaction. Each simulation moves one profile from clean to the given scandal, preserving the observed distribution of alignment and positive-valence attributes. Amplification is the additional absolute penalty relative to the inter-party prediction; all eight amplification intervals exclude zero.

The type-specific model predicts the largest amplification for the most severe scandals, although the reduced-form amplification estimates in Section 6.4 are not monotonic in severity. Every type predicts nearly identical out-group and in-group intra-party penalties: type-specific severity does not generate the observed in-group surcharge, reinforcing the decomposition in Section 7.

Table F3: Profile-level scandal penalties across four partisan contexts, by scandal type.

Panel A: Scandal penalties

Scandal type	Co-partisan inter-party	Out-partisan inter-party	In-group intra-party	Out-group intra-party
Falsified credentials	−0.117 [−0.149, −0.085]	−0.163 [−0.203, −0.122]	−0.269 [−0.313, −0.225]	−0.239 [−0.283, −0.195]
Illegal donations	−0.146 [−0.181, −0.111]	−0.174 [−0.216, −0.132]	−0.342 [−0.385, −0.299]	−0.311 [−0.355, −0.267]
Sexual misconduct	−0.209 [−0.243, −0.175]	−0.211 [−0.250, −0.173]	−0.424 [−0.469, −0.379]	−0.344 [−0.390, −0.298]
Violent protest while underage	−0.099 [−0.130, −0.067]	−0.122 [−0.163, −0.082]	−0.237 [−0.282, −0.191]	−0.216 [−0.262, −0.170]

Panel B: Direct theoretical contrasts

Scandal type	In-group intra-party − co-partisan inter-party	Out-group intra-party − out-partisan inter-party
Falsified credentials	−0.152 [−0.206, −0.098]	−0.076 [−0.136, −0.017]
Illegal donations	−0.196 [−0.250, −0.142]	−0.137 [−0.197, −0.077]
Sexual misconduct	−0.215 [−0.269, −0.161]	−0.132 [−0.191, −0.074]
Violent protest while underage	−0.138 [−0.193, −0.083]	−0.093 [−0.153, −0.034]

Note: Each entry reports the estimated profile-level vote penalty for the indicated scandal relative to no scandal, followed by its 95% confidence interval in brackets. The sample is restricted to Democratic and Republican identifiers. The joint model interacts scandal type with the same respondent-centered four-context coding used in Table 2, includes the main candidate controls, respondent controls, and task-order fixed effects, and clusters standard errors by respondent. Panel B reports covariance-based contrasts from that model; negative estimates indicate a larger penalty in the intra-party context.

F.3 The threshold condition

Proposition 1(i) states the one comparison the model makes conditional: punishment of the voter's *own* co-partisan rises when the contest moves inside the party only if $\beta > q_s(r)$, so that a scandalized co-partisan may remain viable against an out-partisan but lose once the alternative is also a co-partisan. Table F4 evaluates this condition by reporting $\beta - q_s$ for different scandal types. A positive difference means that party utility is larger than the estimated scandal cost; a negative difference means that the scandal is large enough to offset one party match.

To keep this comparison on a common scale, every row in Table F4 comes from the same type-specific version of Equation (9). This model estimates one party coefficient, $\beta = 1.497$, jointly with a separate cost q_s for each scandal; the same β therefore appears in all four comparisons. Each $\beta - q_s$ is a within-model coefficient contrast, and its standard

error uses the respondent-clustered covariance between the jointly estimated β and q_s, which incorporates uncertainty in both quantities.

The pooled specification in Table 3 instead estimates $\beta = 1.449$ alongside one common any-scandal cost. All coefficients are re-estimated when the common scandal and positive-cue indicators are replaced by type-specific indicators, so the two specifications produce slightly different estimates of the same party-utility parameter. This is a consequence of estimating a nonlinear choice model under different restrictions, not evidence that party utility varies across scandals. Combining β from the pooled model with the q_s estimates from the type-specific model would mix coefficients from separate logit normalizations and would not provide the covariance needed for valid uncertainty estimates. We therefore keep the pooled calibration in Table 3 and use the type-specific model alone in Table F4.

Table F4: The type-specific threshold condition $\beta > q_s$.

Scandal	β	q_s	$\beta - q_s$	95% CI
Falsified credentials	1.497	1.097	0.400	[0.249, 0.551]
Illegal funding / corruption	1.497	1.402	0.095	[−0.059, 0.249]
Sexual misconduct	1.497	1.630	−0.134	[−0.287, 0.020]
Violent protest while underage	1.497	0.887	0.610	[0.464, 0.755]

Note: Partisan respondents only. Positive values of $\beta - q_s$ indicate that the estimated party-alignment utility exceeds the scandal cost; negative values indicate that the scandal cost exceeds party-alignment utility. All rows come from one additive type model that estimates a common party-alignment coefficient and a separate utility cost for each scandal. The repeated value of β is therefore one jointly estimated coefficient, not a separate estimate by scandal type. Confidence intervals account for the covariance between β and each q_s. With cue interactions the exact condition is $\beta > q_s(r)$; Appendix F reports the type-by-cue results.

For sexual misconduct, the 95% confidence interval for $\beta - q_s$ narrowly includes zero, so the data do not clearly distinguish its cost from the value of one party match. Its point estimate nevertheless illustrates the model's severity implication: a scandal that approaches or exceeds the value of party can induce cross-party punishment and therefore retain more of its electoral effect in inter-party contests; moving the choice within party then adds less to its penalty.

F.4 Estimation notes

The candidate-difference binary logit and the candidate-long conditional logit return numerically identical coefficients (differences below 1.5×10^{-12}); all standard errors are clustered by respondent, and MRS confidence intervals use the delta method. In the partisan-only strict model, the implied penalties are -0.165 (inter-party), -0.260 (out-group intra-party), and -0.258 (in-group intra-party): excluding Independents reduces the predicted inter-party penalty while leaving intra-party penalties essentially unchanged, widening the implied separation, and does not generate the extra in-group penalty. Under strict and hybrid codings, the flexible $v(s, r)$ models show that credentials and violent protest satisfy the point-estimate condition $\beta > q_s(r)$ in all five cue states, corruption in three of five, and sexual misconduct in none.

F.5 Context-specific logit scale

The common-scale model leaves a larger residual for in-group intra-party contests than for out-group intra-party contests. Because logit coefficients are identified relative to the error scale, lower choice noise in own-party contests could mechanically sharpen the predicted penalty without requiring a distinct scandal cost. We therefore estimate a heteroskedastic-logit check in which the error scale varies by contest context, normalizing the inter-party scale to 1.

Table F5 reports the main quantities. The three-context heteroskedastic model estimates an out-group intra-party scale of 0.887 and an in-group intra-party scale of 0.733. The resulting predicted penalties are -0.172 in inter-party contests, -0.258 in out-group intra-party contests, and -0.296 in in-group intra-party contests, compared with the reported reduced-form penalties of -0.187, -0.275, and -0.310. The in-group residual falls from about 5.2 points under the common-scale model to about 1.4 points under the heteroskedastic model.

This check should be read cautiously. Scale differences are not themselves a behavioral

Table F5: Context-specific logit scale and model-implied scandal penalties.

Contest context	Scale	Common-scale pred.	Heteroskedastic pred.	Reduced-form
Inter-party	1.000	−0.192	−0.172	−0.187
Out-group intra-party	0.887	−0.259	−0.258	−0.275
In-group intra-party	0.733	−0.257	−0.296	−0.310

Note: The heteroskedastic logit normalizes the inter-party scale to 1 and estimates separate scales for out-group and in-group intra-party contests. Predicted penalties simulate a clean-to-scandal change for one profile, preserving the observed distribution of alignment and positive-valence attributes. Reduced-form penalties are the estimates from Section 6.1.

mechanism, and the model does not identify whether lower in-group scale reflects attention, betrayal, moral expectations, or something else. The result is therefore best read as a sensitivity check: much of the extra in-group residual may be mechanical, while the main substitutability result does not depend on that residual.