Can Politicians Say That?
What Shapes Public Responses to Speech Scandals

Conditional Acceptance at: *Journal of Experimental Political Science*

Abstract

Why do some politicians face greater backlash for using insensitive language against identity groups while others do not? Existing explanations focus either on the content of speech or the context in which it occurs. In this article, we propose an integrated framework that considers both and test it using a preregistered conjoint survey on a national U.S. sample. Our findings provide partial support for our expectations. Subjects react most negatively to insensitive speech when the target belongs to their own identity group, when aggravating circumstances exist, and when politicians are of an opposing political party. Our article extends growing scholarship on speech scandals, which has largely explained the fates of politicians as a function of a small number of causative variables in isolation.

Keywords: speech, language, identity, political behavior, experiment

Politicians, to put it mildly, often stick their feet in their mouths. Some live to fight another day, and some never recover. In 2006, former Virginia Sen. George Allen was caught on camera calling a videographer a "macaca" (monkey) (Tapper & Kulman, 2006), and he lost his seat in a fit of disgrace. In 2016, Donald Trump was exposed on a hot mic from a decade earlier bragging that he could "grab [women] by the pussy" (Times, 2016), and he was elected president. There is no shortage of high-profile politicians who come under fire for making insensitive remarks. Amid hotly-charged debates over "cancellations," these controversies matter. What can and cannot be said not only implicates politicians. It has become one of the most contested terrains over which America's "culture wars" are fought (Clark, 2020; Friedersdorf, 2022; Mishan, 2020). Even a cursory look at anecdotes, however, suggests that the consequences for violating speech norms can vary markedly. Why?

In popular circles, answers tend to highlight the remarks themselves and how politicians exercise damage control. Yet beyond the precise language that politicians use, observers commonly speculate that not all politicians, at all times, are held to the same standards. Pundits, for instance, have routinely commented on Trump's "teflon" reputation in which nothing he says—from trafficking in racism to demeaning "locker room talk" against women—seems to "stick" to him (Cillizza, 2016). One *CNN* reporter has written about a "political double standard for ugly speech" (Sheffield, 2018). Such examples are not confined to the U.S. In the U.K., a writer for a major British newspaper argued that a conservative MP "fe[lt] she c[ould] say whatever she wants, because she's a woman and she's queer, and therefore in her eyes she's not like those other party members" (Necati, 2017).

What such counterfactuals have in common is a simple intuition: that the words politicians use may not be the only (or even most salient) factor determining their fates. A growing body of political science research has looked at how the content of speech—ranging from inflammatory racial rhetoric to sexist and other tropes—can affect voter responses to politicians (Benoit, 2017; Conway et al., 2017; Hodges, 2020; Ott, 2016; Newman et al., 2021; Rhodes et al., 2020; Shafer, 2017). At the same time, a parallel scholarship has examined how any array of situational factors—including

the conditions under which words are spoken and details about a speaker—impact public reactions (Christiani, 2023; Lindner & Nosek, 2009; McGraw, 1990; Thompson & Busby, 2023). To date, however, research has not offered an integrated account of how such variables shape the outcomes of speech scandals.

This gap represents a challenge. Despite efforts to conceptualize other serious types of politician misconduct (Dewberry, 2015; Dziuda & Howell, 2021; Invernizzi, 2016), there has been no similarly comprehensive attempt to explain divergent outcomes of speech acts. Even if differences in the controversial words politicians use or the situations in which they say them "matter," the relative importance of these variables is unclear. Voters weigh many factors when evaluating candidates, with some taking precedence. When examined in isolation, individual features of a speech controversy may appear more salient than they in fact are (De la Cuesta et al., 2022; Horiuchi et al., 2020; Rabinowitz et al., 1982). How voters prioritize different criteria—including both the content of speech and the situation surrounding it—requires direct comparisons.

This article fills that gap by developing a general framework to explain public reactions to inappropriate language by politicians. We focus on identity-based speech because of its highly fraught nature. The diffusion of "social justice" movements, including Black Lives Matter, #MeToo, LGBTQ+ advocacy, and the Israel-Palestine issue has placed identity-based speech squarely in the public eye (Bernstein, 2005; Fukuyama, 2018; Lilla, 2018). In doing so, it has galvanized members of the public either advocating for more sensitivity or claiming that the "Overton window" for incorrect speech has narrowed too far (Costello et al., 2019; Haidt & Lukianoff, 2018; Howard, 2019). The features of speech transgressions that we highlight cover the basic "who, what, where, and when" of controversies that would likely feature in media coverage (Tumber & Waisbord, 2019).

To test our predictions, we fielded a preregistered, national online conjoint experiment in the U.S. The experiment asked respondents to evaluate fictitious news articles about politicians, where we randomized key features of the content of the controversial remark and the situation in which it occurred. Our findings provide partial support

for our expectations: As predicted, subjects react most negatively to insensitive speech when the target belongs to their own identity group, when aggravating circumstances exist, and when politicians are of an opposing political party. However, contrary to predictions, subjects do not respond most negatively to slurs, react similarly regardless of how a politician addresses the scandal after the fact, and are no more likely to rule out voting for a politician based on having dissimilar demographic traits.

Theoretically, our study can help to explain why only some politicians have their careers derailed (or even ended) by making derogatory comments. Whereas prior work on speech controversies has focused on a small number of causative factors in isolation, we offer a generalized framework. Empirically, we provide—to our knowledge—the first exploratory test of how a large set of contentual and situational variables compare in determining the outcomes of politicians accused of "wrongspeak." Broadly, our study adds to a considerable literature on scandals in American politics, which has focused on topics like sex allegations, financial misconduct, and other malfeasance (Barnes et al., 2020; Basinger & Rottinghaus, 2012; Busby, 2022; Bowler & Karp, 2004; Doherty et al., 2014; Funk, 1996; Green et al., 2018; Knight, 2011; Maier, 2010; Vonnahme, 2014).

How voters react to speech controversies

We propose a framework in which both the content of the speech and situation in which it happens can shape public judgments of speech scandals. By content, we refer to: 1) the nature of the remarks (what was said and who was targeted); and 2) the politician's response (whether the candidate apologizes, makes excuses, defends the speech, or offers "no comment"). By situational factors, we refer to: 1) the context (the timeframe, the degree of spontaneity, and whether the words reflect a pattern); and 2) the politician's background traits (his or her political party, race, gender, and age). Our pre-registered predictions are as follows:

Contentual factors

1) Nature of original remarks

Reporting on identity-based speech controversies starts with details of the speech itself (Ayo et al., 2020; Kennedy et al., 2018). More severe comments should generally generate greater backlash. Relative to other forms of insensitive speech—including dehumanizing language, stereotypes, and denials of discrimination—slurs, which impugn a group's innate self-worth, tend to be regarded as most transgressive (Croom, 2013; McWhorter, 2021). As such, the public may be most likely to condemn their use as a "cardinal sin." Separately, in line with social identity theory (Ellemers & Haslam, 2012; Tajfel & Turner, 2004), which presumes that people feel more commonality with in-group members, voters may be especially activated if they are members of the same identity group being targeted. These vicarious feelings accord with the "homophily" principle rooted in in-group preferences (McPherson et al., 2001).

- **Prediction 1a:** *Voters will object most to insensitive speech when a slur is used.*
- **Prediction 1b:** *Voters will object most to insensitive speech when their own identity group is "congruent" with the target.*

2) Politician's response

Politicians routinely try to control public relations after a speech scandal. Studies largely suggest that voters will be most punitive toward politicians who expressly try to shift blame away from themselves (Chanley et al., 1994; Kitagawa & Chu, 2021; Thompson & Busby, 2023). Compared to its main alternatives—apologizing or defending one's words—making an excuse for inappropriate language has been shown to be ineffective. Whereas apologies, defenses, and even "no comments" avoid blameshifting, excuses deflect culpability. Research shows that standard excuses—such as pleading ignorance or claiming to misspeak or to be quoted out of context—often backfire by making public officials seem weak or unwilling to "take the heat" (McGraw, 1990). This accords with analyses of groups being more positively inclined toward leaders who insist that "the buck stops" with them (Miller & Reeves, 2022).

- **Prediction 2:** *Voters will object most to insensitive speech when politicians make excuses for their words.*

Situational factors

1) Context

Context may serve an aggravating factor in assigning guilt (Nyhan, 2014, 2017; Schein, 2020). Several issues may be relevant. First, the public may be more critical of biased comments that are more recent (Schulte, 2021). Because definitions of prejudice tend to become more capacious over time (Greenland et al., 2022), with an informal "statute of limitations" attached to past mistakes, voters may be reluctant to apply today's standards of morality to dated words. Planned comments (e.g., at a speech or on social media) might also offer politicians less latitude to eschew accountability or to clarify their remarks than ones delivered more spontaneously (e.g., in an interview, Q&A, or debate) (Frantzich, 2012). Finally, repeated usage of controversial language may make it harder for voters to discount statements as an anomaly or "out of character" (Agadjanian et al., 2019; Nisan & Horenczyk, 1990).

- **Prediction 3:** *Voters will object most to insensitive speech when there are more aggravating factors to impute guilt (that is, when remarks are more recent, planned, and part of a pattern of behavior).*

2) Politician's background traits

Voters often turn to politician background traits as "shortcuts," or heuristics, to make inferences (Campbell & Cowley, 2014; McGraw, 1990). One way is by "projecting" on politicians certain attitudes or biases. When voters are negatively predisposed toward a politician, they may evaluate bad behavior more harshly. Evidence indicates that this is especially likely if their own background characteristics do not match those of the accused politician. Rejection may be conditioned by either the personal political attachments of voters (i.e., party identification) or their ascriptive traits (i.e., race, gender, and age). Extensive research, for example, shows that voters empathize less

with candidates who are on opposing political "teams" (Green et al., 2002) and who do not "look like" them (Aichholzer & Willmann, 2020). Affective disalignment may even trump more programmatic or policy concerns when voting.

- **Prediction 4:** *Voters will object most to insensitive speech when their partisanship and ascriptive traits (race, gender, and age) are "incongruent" with the politician.*

Survey and data collection

Appendix Table A.1 summarizes the full set of experimental variables and hypotheses. To test our predictions, we fielded a preregistered online survey experiment via YouGov U.S. between December 5th and 19th, 2024.[1] The conjoint design randomly manipulated core elements of speech allegations (nature of the remarks, politician's response, context, and politician's background traits) to discern public reactions. Respondents included 3,162 adults drawn from a national panel, weighted to approximate the U.S. population by age, gender, race, education, region, income level, and political party. Each respondent received modest compensation for participating in our study, equivalent to between $5 and $7.5. Due to the sensitivity of the topic, all respondents were warned prior to the experiment that the content could be offensive, with the choice given to opt out. Subjects were debriefed that the scenarios were fictitious. One scope condition on our study is that, by necessity, we assumed that everyone in the public is aware of the speech scandal, which may be more or less realistic depending on actual media coverage and voter attentiveness.

Conjoint design

For each vignette, respondents evaluated in isolation a single candidate portrayed as running for state legislature in a general election who had made a controversial remark, then repeated this task three times (for a total N of 10,200). Conjoints maximize statistical efficiency by measuring the impact of randomly isolated variables on voter

[1] Our experiment was approved by the IE University research committee as being in compliance with ethics policies and guidelines on November 25th, 2024, with reference IERC-16/2024-2025. It was preregistered at OSF, with protocol available from `osf.io/78jf5`.

preferences (Bansak et al., 2021b; Hainmueller et al., 2014). The design addresses the likelihood that making inappropriate remarks is collinear with other traits of politicians that may depress ballot box appeal. It also precludes assessments of the often idiosyncratic nature of real-world speech controversies. We presented just one candidate for respondents to evaluate at a time because it is more realistic for the setting. Although forced-choice conjoints involving candidate "face offs" are common, it is rare that voters would be presented in an actual election with a choice of two politicians, both of whom had made insensitive comments.

Respondents read about the politician and the speech allegations in a brief vignette, similar to a write-up in a newspaper.[2] In total, we randomized ten independent variables. The dedication required to the survey is below the likely level that would trigger "satisficing" behavior. Research shows that respondents are capable of processing a relatively large number of dimensions while maintaining the integrity of causal effects (Bansak et al., 2018, 2021a). To increase realism, a headshot of the candidate manipulated politician background traits of race (via skin tone and hair color), gender, and age. Alongside listing a party affiliation, we also included racialized, gendered names for each politician.[3] In total, we used 16 pictures, one for each race-gender-age category (see Figure A.1).[4] Figure 1 shows an example vignette.[5]

[2]We presented all the respondents with the following introduction before the vignette: "Below is a newspaper article about a politician running for state legislature in a general election. Please read the article below and answer the following questions."

[3]To maintain consistency of features between younger and older politicians of the same race-gender categories, we altered the age of politicians via the photoshop program.

[4]Duplicate pictures across tasks were avoided by the survey administrator.

[5]To enhance clarity, we made slight revisions to the vignette from our pre-registered plan that did not substantively change the information presented to respondents.

Figure 1: Example of candidate presentation

Politician Makes Insensitive Remark
Candidate under fire for controversial comments
By staff writers

Juan García (pictured below), running in the general election for state legislature is facing backlash over controversial comments.

The remarks involved a slur targeted against blacks: "It's hard, I have to deal with lots of n**gers."

The politician made the comments yesterday. They were part of a unplanned conversation. The remarks were made in private.

The candidate has been accused of making similar comments in the past.

The politician responded, "I'm sorry for my remark. It was hurtful, and I take full responsibility."

The remarks have generated controversy surrounding the Democrat politician, who recently won the primary.

We took particular care in representing the content of speech. Media outlets follow different style guidelines in reporting on insensitive language.[6] Apart from a "clean" control that only describes the politician and does not reference any speech scandal, we presented realistic quotes that denoted: 1) slurs; 2) stereotypes; 3) dehumanizing language; or 4) denials of discrimination. This precluded respondents needing to make assumptions about what these generic terms mean. Slurs[7] and stereotypes were necessarily specific to each identity group. To ensure that responses were not driven by idiosyncrasies in the chosen language (Bertrand & Mullainathan, 2004; Sen & Wasow, 2016), we randomized two quotes each (eight in total) varying the precise language for each treatment, then collapsed them into the four central categories (see Table A.2). This also reduces the concern that some comments directed against the

[6]For instance, the *New York Times* associate managing editor for standards reports that "[t]he Times, like many news organizations, generally avoids repeating racial slurs, obscenities and other potentially offensive language" (Corbett, 2015). Other organizations do print such language, often with a warning that an article contains offensive language.
[7]Written with asterisks.

main identity categories (e.g., specific stereotypes for different groups) are regarded as more derogatory.

For the politician's response, we used quotes that represented three different variants of apologies, excuses, and defenses that are commonly invoked. Apologies included: 1) a remorseful apology that expressed simple contrition; 2) a "woke" apology that admitted complicity in broader power structures; or 3) a "sorry-I-offended" apology that conveyed regret that the statement hurt feelings. Excuses included: 1) pleading ignorance; 2) claiming to misspeak; or 3) maintaining that remarks were taken out of context. Defenses included: 1) denying wrongdoing; 2) playing the victim; or 3) going on the attack.[8] Collectively, we randomized nine independent treatments (see Table A.3) equating to the response sub-types, in addition to "no comment" controls. For each of the conditions, we again used two quotes each (20 in total) to reduce idiosyncratic reactions to particular language, and then combined them into the four main categories.

We included three main dependent variables (see Appendix - DV Questions). The first was a simple binary question measuring electoral preferences: "If you had to make a choice without knowing more, would you ever consider voting for this candidate?" (Yes/No). We categorize a "No" answer as the respondent "ruling out" voting for the candidate. The second question asked respondents to evaluate the politician on a "feeling thermometer," based on a standard measure from the American National Election Study. The continuum ran on a 0-100 scale from "coldness" to "warmth." Finally, we asked a general question about the speech act itself: "On a scale of 1 to 7, how objectionable do you think the candidate's behavior is in this case? (From 1 ("Not at all objectionable") to 7 ("The most objectionable I can imagine")). We framed the question in the negative because we assumed that most respondents would reduce their support for politicians accused of insensitive statements.

[8] As with the politician headshots, we requested that YouGov avoid showing duplicate quotes across tasks.

Data and empirics

Power calculations

We conducted a power calculation using the procedure in Schuessler & Freitag (2020) where we sought to identify minimum detectable effects (MDEs) as small as 0.05. To do so, we relied on results from a pilot study with a convenience sample of volunteers (N=919) that we fielded via the Harvard Digital Lab for the Social Sciences (DLABSS) from Dec. 2023 - Jan. 2024.[9] Using estimated coefficients from the pilot on a main "ruling out" binary DV, we calculated (for our maximum number of levels (5) across our attributes, including interactions), a required N of 9,394 (accounting for the loss of power due to including a clean control with no reference to a speech controversy, which lacks variation for the remaining attributes). This was less than our final sample size of 9,486 observations (three tasks per individual respondent). When adjusting our estimates using the procedure in Storey & Tibshirani (2003) to account for testing several hypotheses with effects expected to be of different sizes, we found that we should still be able to detect, for most scenarios, significance (at q <.05) except in the scenario when coefficients are tiny (less than .001, or 1/40th of the smallest coefficient in our pilot).

Data pre-processing

YouGov data are generally high-quality and have been used in a range of recent American politics studies on public opinion (e.g. Bartels, 2018; Huddy et al., 2015). As part of data pre-processing, we took standard checks to ensure the approximation of national demographic representativeness and balanced randomization (see Table A.5, which shows some differences with the population typical of online survey samples). This justifies the use of respondent covariates in our models, as described below.

[9] See anonymized preregistered experiment for the pilot in https://osf.io/ds658/?view_only=3e2c6e0a92304e65bf4d51bcbf46abb1.

Analysis and Results

For our analyses, we estimate marginal means (MMs) under each randomized condition using simple linear models with each level of the attribute as a dummy. MMs have increasingly become standard in conjoint research because they allow for substantive interpretations of the baseline estimands and are useful for comparing the relevance of attributes across subgroups or interactions given different baselines (Abramson et al., 2022; Leeper et al., 2020). We also report the statistical significance of main effects using differences in Average Marginal Component Effects (AMCEs) with q-values, computed in accordance with the Storey-Tibshirani method. These are found to have the best trade-off between excessive Type I and Type II errors under multiple hypothesis testing. Covariates are used in all the models (respondent gender, age, race, education, region, income level, and partisanship).[10]

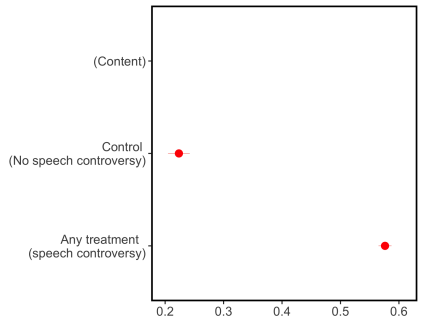
Before turning to our main models, we first estimate a baseline regression, the marginal means of which are shown in Figure 2 Panel A, that compares responses to the "clean" control to those of any treatment condition that references a speech scandal. This estimation indicates the net effect of insensitive remarks on a respondent's likelihood of "ruling out" voting for a candidate accused of controversial remarks. As expected, respondents are significantly more likely to rule out voting for politicians associated with a speech controversy, more than doubling the probability that they would never consider voting for the candidate.

In Panels B-F of Figure 2, we also check whether some respondents react more negatively in the aggregate to insensitive speech by stratifying respondents by political party/ideology and demographics (race, gender, and age). In terms of party identification and ideology, research suggests that Democrats and liberals more generally may be more sensitive to concerns about harm to vulnerable groups (Graham et al., 2009). In line with this prediction, we find that both Democrats (compared to Republicans) (Panel B) and liberals (compared to moderates and conservatives, re-

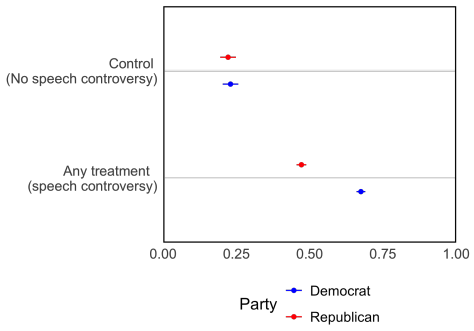
[10] Gender defined as man or woman; age defined as older (60+) or part of the Millennial generation or younger (born 1981 or after); race defined as white or non-white; education defined as college-graduate or not; region operationalized as fixed effects by U.S. Census-designated areas (Northeast, Midwest, South, and West); income measured as a continuous variable; partisanship defined as Democrat, Republican, or other.

spectively) (Panel C) are significantly more likely to rule out voting for politicians who have made insensitive remarks. Regarding demographics, scholarship indicates that aversion to insensitivity may be more pronounced among racial minorities (Hutchings & Valentino, 2004), women (Bittner & Goodyear-Grant, 2017), and youth voters (Holbein & Hillygus, 2020), for whom identity politics is potentially more salient. As expected, with gender, we find that female respondents are more likely than male respondents to rule out voting for politicians linked to controversial speech (Panel E). However, results yield no clear statistical differences for respondents of different races (Panel D) or age cohorts (Panel F).

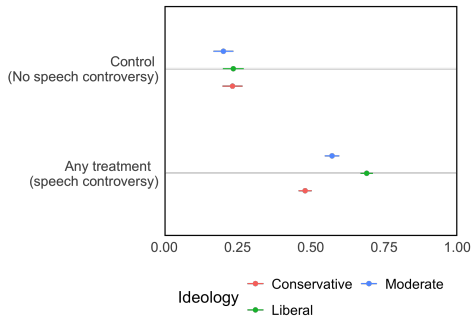
Figure 2: Marginal means (MMs) of the probability of answering "I would NEVER consider voting for this candidate," contrasting respondents who receive no information about controversial speech by the candidate with those who are informed that the candidate has been involved in a speech scandal



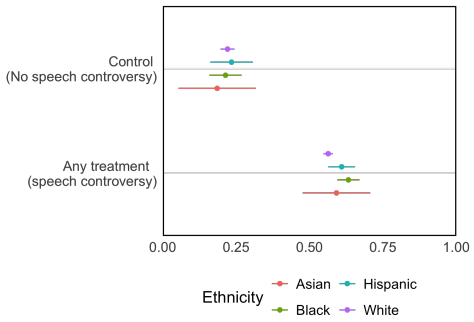
(A) Whole sample



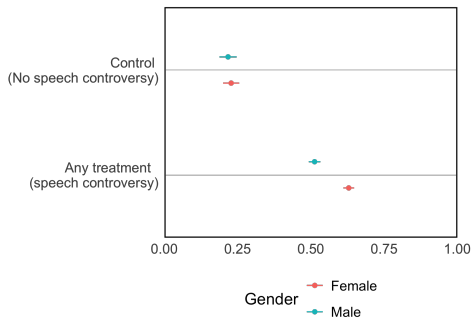
(B) By party



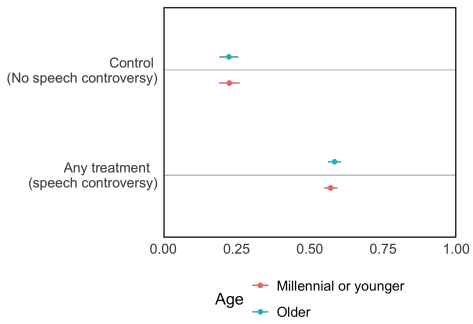
(C) By ideology



(D) By race



(E) By gender



(F) By age

Note: Within models, all treatments are statistically significant, calculated as an AMCE (q <.05).

For our main models, we compare all experimental conditions within their respective treatment categories. As explained above, we made specific predictions only about which one variable within each category leads to the highest voter rejection of politicians. We fit linear models of the form:

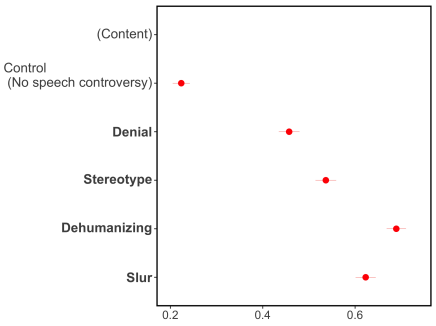
$$Y_{ik} = \sum_{j} \beta_j A_{jik} + X_i + \epsilon_i \tag{1}$$

where Y is the DV (binary of ruling out voting for the candidate, thermometer scale of "warmth" toward the candidate, and scale of objection to behavior), A is a vector of dummy variables for each attribute level j (our IVs), and X is the series of respondent-level covariates. We binarize the "aggravating factors" variables and sum them together for a total value (from 0 to 3). We also binarize the "respondent-politician incongruence" variables based on whether respondents have all incongruent background traits with politicians or not. Our respondents are indexed by i, and k (1-3) indexes for one of the three candidates that respondents assess. Our main IVs are summarized in Table A.6.

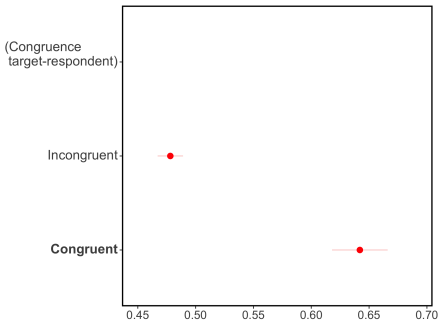
Main estimations. Figure 3 displays the MMs associated with the main DV of ruling out voting for the candidate. Alongside these estimations, Figure A.2 reports AMCEs with confidence intervals.

Panel A reveals that, contrary to Prediction 1a, respondents react most negatively to politicians accused of using dehumanizing language, not slurs. However, slurs are rated more objectionable than stereotyping or denying that discrimination exists against identity groups. In Panel B, consistent with Prediction 1b, respondents are significantly more likely to rule out voting for politicians when the target of the insensitive speech belongs to their own identity group. Panel C shows that, contrary to Prediction 2, which anticipated that excuses would elicit the most negative reaction, no specific type of politician response statistically stands out as the most likely to increase the likelihood of respondents ruling out voting for politicians. In fact, politicians who make excuses actually tend to fare better than those who defend their insensitive speech or (perhaps surprisingly) those who offer no comment. Panel D supports Prediction 3, indicating that the presence of more aggravating factors in-

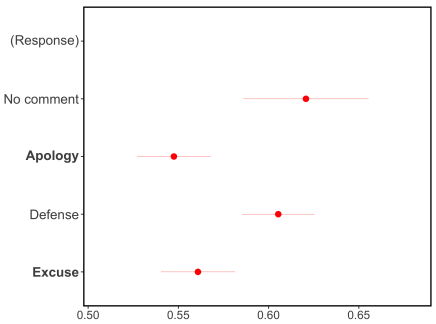
Figure 3: Marginal means (MMs) of the probability of answering "I would never consider voting for this candidate"



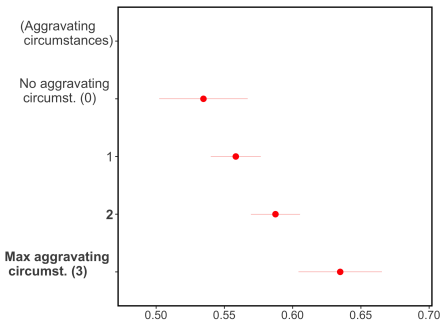
(A) Pred. 1a: Nature of original remarks (severity)



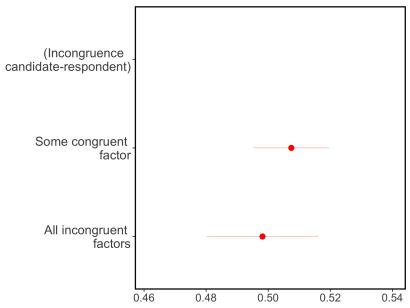
(B) Pred. 1b: Nature of original remarks (target) (congruence with respondent)



(C) Pred. 2: Politician's response



(D) Pred. 3: Context



(E) Pred. 4: Politician's background traits (incongruence with respondent)

Note: In panels A-E, we **boldface** variables for which the difference with the reference category is statistically significant, calculated as an AMCE (q <.05). The reference category is the topline variable in each case ("no speech controversy," "incongruent," "no comment" "no aggravating circumstances," "some congruent factor"). Full AMCE results are shown in Figure A.2, with point estimates, t-statistics and q-values for all variables in Table A.7.

creases the likelihood of respondents being unwilling to vote for politicians. Finally, Panel E is inconsistent with Prediction 4, showing that respondents are no more or less likely to rule out voting for politicians when all of the respondent's background traits are incongruent with the politician's. Figure A.3 shows the AMCEs for the two alternative DVs: the 1-100 feeling thermometer, and 1-7 scale of how objectionable the respondent finds the candidate's behavior. In both cases, results are substantively identical to those of our main "ruling out" variable.

Additional analyses. In addition to estimating aggregate subgroup effects as described above, Figures A.4-A.8 also report our main findings with a breakdown of respondents by political party/ideology and demographics (race, gender, and age). Directionally, results are similar to the main models for the various subgroups, although less precisely estimated due to the smaller sample size within each category. Figure A.9 unpacks the effects of respondent congruence with the target of a speech controversy (by race, religion, sexual orientation, and gender). All of the effects are directionally positive as expected and statistically significant, showing that any form of respondent-target congruence makes subjects more likely to rule out voting for politicians. Figure A.10 disaggregates the respective effects of aggravating factors that we collapsed into a composite. Results are again directionally positive as predicted and statistically significant for each factor. Respondents are more apt to rule out voting for politicians whose statements are more recent, planned, and reflect a pattern. Lastly, Figure A.11 disaggregates the effects of respondent incongruence with a politician's background traits by political party and demographics (race, gender, and age). All of the effects are directionally positive as predicted, though only political party is statistically significant. Respondents are more inclined to rule out voting for politicians accused of insensitive speech who are of the opposing party.

Robustness checks

We conduct several robustness checks to ensure the reliability of our findings. First, given that we present candidates running for state legislature, and models average over the demographics of candidates presented (De la Cuesta et al., 2022), Figure A.12

Panel A weights observations so that the demographics of the candidates in the experiment (by race, gender, and age) match those of elected officials in state legislatures across the U.S.[11] Additionally, because we do not observe the full target distribution of controversial speech acts, we also use a procedure analogous to De la Cuesta et al. (2022) in Figure A.12 Panels B and C that "bounds" our estimates. Specifically, we reweight our sample under a "low expected rule out" scenario (where observations denoting attribute levels expected to most correlate with voters not ruling out voting for a politician are double-weighted) and a "high expected rule out" scenario (where observations denoting attribute levels expected to most correlate with voters ruling out voting for a politician are double-weighted). In general, results look substantively similar as our main findings. Lastly, as described above, we test whether the results are affected by idiosyncracies in the phrasings of the controversial remarks by the politicians or their responses. Figure A.13 estimates differentiated effects for quotes under each category, as presented in Table A.2 and Table A.3. Results again appear substantively robust to alternative wordings of the politician quotes. However, it is worth mentioning that we do find considerable variability in different kinds of insensitive speech depending on both the group targeted and the particular language used. For example, in multiple instances, the use of slurs had similar effects as dehumanizing language, even though the former produced larger negative reactions overall. In some cases, such as slurs targeted at Blacks and Jews, the effects are actually larger than the average effects of using dehumanizing language. In others, such as slurs targeted at women, Hispanics and gays, effects are generally below the average effects of the use of dehumanizing language.

Conclusion

As sensitivities over inappropriate speech has become more pronounced in American politics, understanding why some candidates are impacted differently by the words they say is important. Our multi-part, preregistered survey experiment in the U.S.

[11] We obtain these data from National Conference of State Legislatures (2020), and weight by state population. As far as we are aware, no demographic data exists on candidates (instead of actual elected officials) in state legislatures.

sheds light on how both the content of speech and the situation surrounding it affects voter reactions. Although prior accounts have identified such variables as potentially impactful, studies have not contrasted their relative salience within an integrated framework. In a national survey experiment, we found that respondents are most likely to reject politicians for insensitive speech when the target is a member of their own identity group (race, religion, sexual orientation, and gender), when there are more aggravating circumstances (such as being repeat offenders), and when politicians are of a different party from the respondent's own.

Our findings advance a template of relevant factors shaping public reactions to speech scandals, permitting further studies to add, revise, or delete criteria as appropriate. Future research might expand on our results in several ways. First, experiments could present vignettes of controversial speech that simulate how voters obtain information in other "real-life" scenarios (e.g., via social media feeds). Additionally, qualitative analyses of real speech controversies might trace how the mechanisms that we identify as important operate in practice by changing recipients' beliefs or emotions. Scholars could also investigate how less tangible traits like personality or charisma affect the ability of politicians to eschew backlash from speech scandals. Moreover, they could explore how controversies affect particular demographic groups differently, including potentially galvanizing the support of some subgroups. Finally, studies could examine speech controversies in other contexts, such as journalism, academia, or entertainment.

Our study complements a large literature on public responses to scandals, which has mostly focused on topics like extramarital affairs, financial corruption, and other indiscretions. Despite the attention given in the media and popular culture to speech controversies, the question of why some politicians seem more or less immune to criticism remains ripe for research. Examining how social norms and "evolving standards of decency" change over time, vary both within and across societies, and lead to different outcomes for politicians are key. Such extensions could help political scientists better comprehend the varied responses to politicians accused of "wrongspeak." These issues are likely to remain relevant given the momentum of social justice movements

and moves to redefine what public figures can (and cannot) say across diverse areas.

References

Abramson, Scott F., Korhan Kocak & Asya Magazinnik. 2022. What do we learn about voter preferences from conjoint experiments? *American Journal of Political Science* 66(4). 1008–1020.

Agadjanian, Alexander et al. 2019. Counting the pinocchios: The effect of summary fact-checking data on perceived accuracy and favorability of politicians. *Research & Politics* 6(3).

Aichholzer, Julian & Johanna Willmann. 2020. Desired personality traits in politicians: Similar to me but more of a leader. *Journal of Research in Personality* 88. 103990.

Ayo, Femi Emmanuel, Olusegun Folorunso, Friday Thomas Ibharalu & Idowu Ademola Osinuga. 2020. Machine learning techniques for hate speech classification of twitter data: State-of-the-art, future challenges and research directions. *Computer Science Review* 38. 100311.

Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins & Teppei Yamamoto. 2018. The number of choice tasks and survey satisficing in conjoint experiments. *Political Analysis* 26(1). 112–119.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins & Teppei Yamamoto. 2021a. Beyond the breaking point? survey satisficing in conjoint experiments. *Political Science Research and Methods* 9(1). 53–71.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins & Teppei Yamamoto. 2021b. Conjoint survey experiments. In James N. Druckman & Donald P. Green (eds.), *Advances in experimental political science*, Cambridge, UK: Cambridge University Press.

Barnes, Tiffany D, Emily Beaulieu & Gregory W Saxton. 2020. Sex and corruption: How sexism shapes voters' responses to scandal. *Politics, Groups, and Identities* 1.

Bartels, Larry M. 2018. Partisanship in the trump era. *Journal of Politics* 80(4). 1483–1494.

Basinger, Scott & Brandon Rottinghaus. 2012. Skeletons in the white house closets: An empirical investigation into modern presidential scandals. *Political Science Quarterly* 127(2). 213–239.

Benoit, William. 2017. Image repair on the donald trump 'access hollywood' video: 'grab them by the p*ssy'. *Communication Studies* 68(3). 243–259.

Bernstein, Mary. 2005. Identity politics. *Annual Review of Sociology* 31. 47–74.

Bertrand, Marianne & Sendhil Mullainathan. 2004. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review* 94(4). 991–1013.

Bittner, Amanda & Elizabeth Goodyear-Grant. 2017. Digging deeper into the gender gap: Gender salience as a moderating factor in political attitudes. *Canadian Journal of Political Science/Revue Canadienne De Science Politique* 50. 559–578.

BLS. 2024. The employment situation - august 2024. https://www.bls.gov/news.release/pdf/empsit.pdf.

Bowler, Shaun & Jeffrey A Karp. 2004. Politicians, scandals, and trust in government. *Political Behavior* 26(3). 271–287.

Busby, Robert. 2022. *Scandal and american politics in the 21st century*. Switzerland: Palgrave Macmillan.

Campbell, Rosie & Philip Cowley. 2014. What voters want: Reactions to candidate characteristics in a survey experiment. *Political Studies* 62(4). 745–765.

Center, Pew Research. 2024. The partisanship and ideology of american voters - april 2024. https://www.bls.gov/news.release/pdf/empsit.pdf.

Chanley, Virginia, John L Sullivan, Marti Hope Gonzales & Margaret Bull Kovera. 1994. Lust and avarice in politics: Damage control by four politicians accused of wrongdoing (or, politics as usual). *American Politics Quarterly* 22(3). 297–333.

Christiani, Leah. 2023. When are explicit racial appeals accepted? examining the role of racial status threat. *Political Behavior* 45. 103–123.

Cillizza, Chris. 2016. Why does donald trump get away with saying whatever he wants? *Washington Post* https://www.washingtonpost.com/news/the-fix/wp/2016/03/22/how-donald-trump-gets-away-with-saying-all-of-the-controversial-things-he-says/.

Clark, Meredith. 2020. Drag them: A brief etymology of so-called 'cancel culture'. *Communication and the Public* 5(3-4). 88–92.

Conway, Lucian Gideon, Meredith A. Repke & Shannon C. Houck. 2017. Donald trump as a cultural revolt against perceived communication restriction: Priming political correctness norms causes more trump support. *Journal of Social and Political Psychology* 5(1). 244–259.

Corbett, Philip B. 2015. The times's policy on racial slurs: Editor's notebook. *The New York Times* https://archive.nytimes.com/www.nytimes.com/times-insider/2015/06/22/the-timess-policy-on-racial-slurs-editors-notebook/.

Costello, Matthew, James Hawdon, Colin Bernatzky & Kelly Mendes. 2019. Social group identity and perceptions of online hate. *Sociological Inquiry* 89(3). 427–452.

Crabtree, Charles, Jae Yeon Kim, S Michael Gaddis, John B Holbein, Cameron Guage & William W Marx. 2023. Validated names for experimental studies on race and ethnicity. *Scientific Data* 10(1). 130.

Croom, Adam M. 2013. How to do things with slurs: Studies in the way of derogatory words. *Language & Communication* 33(3). 177–204.

De la Cuesta, Brandon, Naoki Egami & Kosuke Imai. 2022. Improving the external validity of conjoint analysis: The essential role of profile distribution. *Political Analysis* 30(1). 19–45. doi:10.1017/pan.2020.40.

Dewberry, David. 2015. *The american political scandal: Free speech, public discourse, and democracy*. New York, NY: Roman and Littlefield.

Doherty, David, Conor M. Dowling & Michael G. Miller. 2014. Does time heal all wounds? sex scandals, tax evasion, and the passage of time. *PS: Political Science & Politics* 47(2). 357–366.

Dziuda, Wioletta & William G. Howell. 2021. Political scandal: A theory. *American Journal of Political Science* 65(1). 197–209.

Ellemers, Naomi & S. Alexander Haslam. 2012. Social identity theory. In Paul A. M. Van Lange, Arie W. Kruglanski & E. Tory Higgins (eds.), *Handbook of theories of social psychology*, vol. 2, 379–398. Los Angeles, CA: Sage.

Frantzich, Stephen. 2012. *O.o.p.s. observing our politicians stumble: The worse candidate gaffes and recoveries in presidential campaigns*. Santa Barbara: Praeger.

Friedersdorf, Conor. 2022. The real reason cancel culture is so contentious. *The Atlantic* .

Fukuyama, Francis. 2018. *Identity: The demand for dignity and the politics of resentment*. New York: Farrar, Straus and Giroux.

Funk, Carolyn L. 1996. The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior* 18(1). 1–24.

Graham, Jesse, Jonathan Haidt & Brian A. Nosek. 2009. Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology* 96(5). 1029–1046.

Green, Donald, Bradley Palmquist & Eric Schickler. 2002. *Partisan hearts and minds: Political parties and the social identity of voters*. New Haven, CT: Yale University Press.

Green, Donald P, Adam Zelizer & David Kirby. 2018. Publicizing scandal: Results from five field experiments. *Quarterly Journal of Political Science* 13(3). 237–261.

Greenland, Katy, Keon West & Colette van Laar. 2022. Definitional boundaries of discrimination: Tools for deciding what constitutes discrimination (and what doesn't). *Journal of Applied Social Psychology* 52(10). 945–964.

Haidt, Jonathan & Greg Lukianoff. 2018. *The coddling of the american mind: How good intentions and bad ideas are setting up a generation for failure*. New York: Penguin.

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

Hodges, Adam. 2020. *When words trump politics: Resisting a hostile regime of language*. Stanford: Stanford University Press.

Holbein, John B. & D. Sunshine Hillygus. 2020. *Making young voters: Converting civic attitudes into civic action*. Cambridge: Cambridge University Press.

Horiuchi, Yusaku, Daniel Smith & Teppei Yamamoto. 2020. Identifying voter preferences for politicians' personal attributes: A conjoint experiment in japan. *Political Science Research and Methods* 8(1). 75–91. doi:10.1017/psrm.2018.26.

Howard, Jeffrey W. 2019. Free speech and hate speech. *Annual Review of Political Science* 22. 93–109.

Huddy, Leonie, Lilliana Mason & Lene Aaroe. 2015. Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review* 109(1). 1–17.

Hutchings, Vincent & Nicholas A. Valentino. 2004. The centrality of race in american politics. *Annual Review of Political Science* 7. 383–408.

Invernizzi, Giovanna. 2016. Political scandals. Columbia University mimeo.

Kennedy, Brendan, Drew Kogon, Kris Coombs, Joseph Hoover, Christina Park, Gwenyth Portillo-Wightman, Aida Mostafazadeh Davani, Mohammad Atari & Morteza Dehghani. 2018. A typology and coding manual for the study of hate-based rhetoric. *PsyArXiv* .

Kitagawa, Risa & Jonathan A. Chu. 2021. The impact of political apologies on public opinion. *World Politics* 73(3). 441–481.

Knight, Richard. 2011. Love means sometimes having to say you're sorry. In Alison Dagnes (ed.), *Sex scandals in american politics: A multidisciplinary approach to the construction and aftermath of contemporary political sex scandals*, 113–133. London: Continuum.

Leeper, Thomas J., Sara B. Hobolt & James Tilley. 2020. Measuring subgroup preferences in conjoint experiments. *Political Analysis* 28(2). 207–221.

Lilla, Mark. 2018. *The once and future liberal: After identity politics*. Oxford: Oxford University Press.

Lindner, Nicole M. & Brian A. Nosek. 2009. Alienable speech: Ideological variations in the application of free-speech principles. *Political Psychology* 30(1). 67–92.

Maier, Jurgen. 2010. The impact of political scandals on political support: An experimental test of two theories. *International Political Science Review* 32(3). 283–302.

McGraw, Kathleen M. 1990. Avoiding blame: An experimental investigation of political excuses and justifications. *British Journal of Political Science* 20(1). 119–131.

McPherson, Miller, Lynn Smith-Lovin & James M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27. 415–444.

McWhorter, John. 2021. Racism punching down? *It Bears Mentioning* https://johnmcwhorter.substack.com/p/racism-punching-down.

Miller, David & Andrew Reeves. 2022. Pass the buck or the buck stops here? the public costs of claiming and deflecting blame in managing crises. *Journal of Public Policy* 42(1). 63–91.

Mishan, Ligaya. 2020. The long and tortured history of cancel culture. *New York Times* https://www.nytimes.com/2020/12/03/t-magazine/cancel-culture-history.html.

National Conference of State Legislatures. 2020. State legislator demographics. Accessed on: 11-13-2023. https://www.ncsl.org/about-state-legislatures/state-legislator-demographics.

Necati, Yas. 2017. Ruth davidson telling the tory party to 'man up' proves that there's no such thing as a liberal conservative. *Independent* https://www.independent.co.uk/voices/conservative-party-conference-ruth-davidson-man-up-nervous-breakdown-sexist-ableist-a7979306.html.

Newman, Benjamin, Jennifer L. Merolla, Sono Shah, Danielle Casarez Lemi, Loren Collingwood & S. Karthick Ramakrishnan. 2021. The trump effect: An experimental investigation of the emboldening effect of racially inflammatory elite communication. *British Journal of Political Science* 51(3). 1138–1159.

Nisan, Mordecai & Gabby Horenczyk. 1990. Moral balance: The effect of prior behaviour on decision in moral conflict. *British Journal of Social Psychology* 29(1). 29–42.

Nyhan, Brendan. 2014. Scandal potential: How political context and news congestion affect the president's vulnerability to media scandal. *British Journal of Political Science* 45(2). 435–466.

Nyhan, Brendan. 2017. Media scandals are political events: How contextual factors affect public controversies over alleged misconduct by u.s. governors. *Political Research Quarterly* 70(1). 223–236.

Ott, Brian L. 2016. The age of twitter: Donald j. trump and the politics of debasement. *Critical Studies in Media Communication* 34(1).

Rabinowitz, George, James W. Prothro & William Jacoby. 1982. Salience as a factor in the impact of issues on candidate evaluation. *The Journal of Politics* 44(1). 41–63. 23 pages.

Rhodes, Jesse H., Elizabeth A. Sharrow, Jill S. Greenlee & Tatishe M. Nteta. 2020. Just locker room talk? explicit sexism and the impact of the access hollywood tape on electoral support for donald trump in 2016. *Political Communication* 37(6).

Schein, Chelsea. 2020. The importance of context in moral judgments. *Perspectives on Psychological Science* 15(2). 207–215.

Schuessler, Julian & Markus Freitag. 2020. Power analysis for conjoint experiments. *SocArXiv* doi:10.31235/osf.io/9yuhp.

Schulte, Gabriela. 2021. Poll: 69 percent say cancel culture unfairly punishes people for past actions, statements. *The Hill* https://thehill.com/hilltv/what-americas-thinking/580747-poll-69-percent-say-cancel-culture-unfairly-punishes-people-for/.

Sen, Maya & Omar Wasow. 2016. Race as a bundle of sticks: Designs that estimate effects of seemingly immutable characteristics. *Annual Review of Political Science* 19. 499–522.

Shafer, Jessica Gantt. 2017. Donald trump's "political incorrectness": Neoliberalism as frontstage racism on social media. *Social Media + Society* 3(3). 1–10.

Sheffield, Carrie. 2018. Let's acknowledge the political double standard for ugly speech. https://www.cnn.com/2018/06/01/opinions/hypocrisy-offensive-speech-opinion-sheffield/index.html.

Storey, John D & Robert Tibshirani. 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* 100(16). 9440–9445.

Tajfel, Henri & John C. Turner. 2004. The social identity theory of intergroup behavior. In John T. Jost & Jim Sidanius (eds.), *Political psychology: Key readings*, 276–293. Psychology Press.

Tapper, Jake & Betsy Kulman. 2006. The macaca heard round the world. https://abcnews.go.com/Nightline/story?id=2322630&page=1.

Thompson, Andrew Ifedapo & Ethan C. Busby. 2023. Defending the dog whistle: The role of justifications in racial messaging. *Political Behavior* 45. 1241–1262.

Times, New York. 2016. Transcript: Donald trump's taped comments about women. *New York Times* https://www.nytimes.com/2016/10/08/us/donald-trump-tape-transcript.html.

Tumber, Howard & Silvio Waisbord (eds.). 2019. *The routledge companion to media and scandal*. London: Routledge.

US Census. 2024. Quick facts united states. https://www.census.gov/quickfacts/fact/table/US/PST045222.

Vonnahme, Beth Miller. 2014. Surviving scandal: An exploration of the immediate and lasting effects of scandal on candidate evaluation. *Social Science Quarterly* 95(5). 1308–1321.