

**Running with the Double Bind: Exploring
Ideological Moderation in Black Women
Candidates**

Jalen Ward

Department of Political Science

University of California, Davis

Qualifying Examination

April 1, 2026

Abstract

Candidates in the United States often moderate their rhetoric after primary elections to appeal to the broader general electorate (Acree et al. 2020). Yet little work examines whether this well-documented pivot is conditioned by candidates' intersecting identities. This project examines whether Black women candidates moderate their ideological campaign rhetoric more than their peers, in response to compounded racialized and gendered expectations of electability and progressivism. Using Twitter data from U.S. House candidates during the 2024 election, I employ a Difference-in-Differences and Regression Discontinuity Design using the primary election date as a breakpoint. The analysis applies zero-shot natural language inference to track ideological shifts in candidate rhetoric. I also examine shifts in issue framing, focusing on whether appeals tied to race, gender, and identity give way to broader policy themes after the primary. Contrary to expectations, Black women candidates do not systematically moderate their rhetoric after primaries, nor do they moderate differently than other identity groups.

Keywords: ideological moderation, Black women candidates, intersectionality, campaign rhetoric, Twitter, 2024 election

Contents

1	Introduction	4
2	Previous Literature	6
2.1	Strategic Moderation in Two-Stage Elections	6
2.2	Race, Gender, and Candidate Evaluations	7
2.3	Black Women’s Unique Strategic Environment	9
3	Theory	11
4	Data	13
4.1	Candidate Metadata	13
4.2	Candidate Text Data	14
5	Research Design	15
5.1	Ideology: Conceptual Definition and Measurement	15
5.2	Measurement Validation	18
5.2.1	Inter-Rater Reliability	18
5.2.2	Validity with Legislative Ideology Scores	19
5.3	Analytical Strategy	20
5.3.1	Difference-in-Differences	20
5.3.2	Regression Discontinuity	21
5.3.3	H3: Identity Appeals Analysis	22
6	Results	23
6.1	H1: Do Black Women Congressional Candidates Moderate After Primaries?	23
6.2	H2: Do Black Women Moderate More Than Other Groups?	25
6.3	H3: Do Black Women Reduce Identity Appeals After Primaries?	29

7 Conclusion	32
8 Limitations	33
9 Future Directions	34

1 Introduction

In 2023, Cherelle Parker made history as Philadelphia’s 100th mayor and its first woman and first Black woman to hold the office. Her path to City Hall, however, required striking ideological shape-shifting over time. As a city council member under then-Mayor Jim Kenney, Parker had built a reputation as a progressive voice, voting against expansions of the police budget and championing criminal justice reform (Nickels 2023). Yet as a mayoral candidate, she repositioned herself as a law-and-order Democrat, going so far as to suggest she would support the use of stop-and-frisk tactics to combat the city’s gun violence crisis — a position that placed her to the right of much of her party’s base (Griswold 2023). Then once elected, Mayor Parker walked a tightrope, carefully reiterating her public safety commitment and clarifying that “there is no place for unconstitutional stop-and-frisk” (WHYY 2023). The shift in positioning and rhetoric was neither accidental nor incidental — it was strategic, and it was electorally successful.

Parker’s transformation illustrates a tension that existing scholarship has not fully resolved. Ever since Downs (1957) established that rational candidates have an incentive to tack toward the center in general elections, scholars have confirmed that this strategic repositioning is a well-documented feature of American electoral politics (Acree et al. 2020; Hummel 2010; Cowburn and Sältzer 2025). Research further shows that candidate characteristics like incumbency, district partisanship, and electoral vulnerability shape the degree to which moderation occurs (Burden 2001; Ansolabehere, Snyder and Stewart 2001). What is not well established, however, are the conditions under which a candidate’s intersecting racial and gender identities shape the nature and magnitude of this pivot. Scholars in race and ethnic politics have demonstrated that Black women candidates navigate a distinct set of stereotypes that are simultaneously racialized and gendered (Lemi and Brown 2019; Chen and Sorensen 2025; Snipes 2025). These stereotypes may systematically distort how voters perceive their ideological positions, even when their actual stances do not warrant such

attributions (Carew 2016; McDermott 1998).

This paper asks a simple yet underexplored question: do Black women candidates moderate their ideological rhetoric more sharply than other candidates after winning their primaries? This question matters for two reasons. First, it extends the strategic moderation literature by examining whether the well-documented primary-to-general pivot varies systematically by candidates' intersecting identities — a dimension that existing work has largely ignored by treating moderation as uniform across all candidates. Much like scholarship within race and ethnic politics that critiques the discipline's tendency to treat race as an exogenous variable or to study political behavior as if racial identity produces homogeneous effects across all actors (Smith 2004; Hochschild and Weaver 2007), this paper insists that the *who* of strategic moderation is as important as the *whether*. Second, the question speaks to a broader normative concern about representation — if Black women face compounded pressures to suppress their rhetorical identities in order to appear viable, this has implications not only for who runs and wins, but for whose interests are ultimately represented in office.

To test this claim, I assembled an original dataset of over 17,000 campaign tweets from 128 Democratic House candidates during the 2024 election cycle. Using each candidate's primary election date as a treatment threshold, I estimate a series of Difference-in-Differences (DiD) and Regression Discontinuity (RD) models comparing pre- and post-primary ideological rhetoric across six identity groups. I also examine shifts in identity-linked appeals to test whether Black women candidates reduce the frequency of race- and gender-coded language after the primary. Contrary to my expectations, I find no statistically significant evidence that Black women moderate their rhetoric more sharply than other candidates after the primary, nor do they uniquely reduce identity-linked appeals — though suggestive patterns and event-time visualizations do reveal slight heterogeneity in rhetorical shifts across identity groups.

2 Previous Literature

2.1 Strategic Moderation in Two-Stage Elections

Classical theories of electoral competition hold that candidates in two-stage election systems face a fundamental strategic dilemma. During the primary, candidates appeal to their partisan base, interest groups, and ideologically motivated “policy demanders” by adopting more extreme positions (Bawn et al. 2012; Aldrich 1983). Once they secure the nomination, Downs (1957) Median Voter Theorem predicts they will tack toward the center to capture undecided voters in the general electorate. Scholars of American elections have long documented this pattern of running to the base before moderating for mass appeal (Brady, Han and Pope 2007; Sides, Tesler and Vavreck 2018; Acree et al. 2020; Hummel 2010). Recent scholarship extending this logic to congressional candidates finds similar evidence. Cowburn and Sältzer (2025) analyzed the Twitter language of nearly 1,000 U.S. House candidates and found clear evidence of post-primary ideological moderation, while Bailey and Reese (2025) show that moderate positioning on social media and campaign websites tends to enhance electoral outcomes on average.

Yet this moderation is neither uniform nor unconditional. Research shows that candidate characteristics like incumbency, district partisanship, and electoral vulnerability shape the degree to which candidates moderate after the primary (Burden 2001; Ansolabehere, Snyder and Stewart 2001). Wittenbrink (2022) finds that moderation patterns differ by party, with Republican candidates showing stronger post-primary pivots than Democrats in 2020. These findings suggest that the primary-to-general shift is better understood as a conditional phenomenon than a universal one, meaning the magnitude of moderation should vary systematically depending on the strategic environment candidates face. What the relevant literature has not examined, however, is whether candidates’ intersecting racial and gender identities constitute a meaningful source of that heterogeneity. If Black women candidates are systematically perceived as more liberal than their actual positions imply — a pattern

well documented in race and ethnic politics scholarship — they may face compounded incentives to over-correct toward the center in ways that other candidates do not.

2.2 Race, Gender, and Candidate Evaluations

Voters do not evaluate candidates in an ideological vacuum. When candidates belong to historically marginalized groups, identity-linked stereotypes tend to serve as informational shortcuts for voters, shaping perceptions of candidates' traits, competence, and ideological positioning before any policy platform comes into view (McDermott 1998). Understanding how these heuristics operate and, more importantly, how they compound across multiple identities is essential for theorizing why Black women candidates may face distinct incentives to moderate their campaign rhetoric after winning primaries.

An established finding in the candidate evaluation literature is that women candidates are systematically stereotyped as more liberal than their male counterparts (Huddy and Terkildsen 1993; King and Matland 2003; Koch 1999) and are seen as more competent on issues related to social welfare, education, and family policy (Schneider and Bos 2016). This gender-ideology stereotype operates even when candidates' actual positions do not support the ascription, functioning as a baseline prior that voters apply in low-information contexts (McDermott 1998).

Black candidates face a partially overlapping set of stereotypes wherein voters consistently perceive Black candidates as more ideologically liberal than objective indicators would suggest, due in part to historical patterns of Black voters being more likely to align themselves with the Democratic Party (Brewer and Stonecash 2001; Mangum 2013; Miller and Schofield 2008). Likewise, Black candidates, despite evidence that Black Democrats are not as liberal as White Democrats, are seen as more liberal by White voters (Lerman and Sadin 2016; McDermott 1998). Black candidates are also seen as more focused on minority-group concerns and social justice issues than on competency domains like foreign policy and economic management (Sigelman et al. 1995; Fulton and Gershon 2018).

Fulton and Gershon (2018) find that these misperceptions of ideological extremity are not benign—they measurably reduce support among White voters and independents, precisely the constituencies that candidates must attract to win competitive general elections.

Black women candidates do not simply inherit the sum of these two stereotype sets. Rather, scholars working in the tradition of intersectionality — the framework introduced by Crenshaw (1989) to analyze how race and gender produce non-additive, mutually constitutive forms of disadvantage — have demonstrated that Black women occupy a distinct evaluative position that cannot be reduced to being Black plus being a woman (Hancock 2007; Brown and Lemi 2021). Stereotypes of Black women candidates draw on historically rooted judgments about professional demeanor, respectability, legitimacy in authority roles, and body politics that have no direct counterpart in the stereotype profiles applied to Black men or White women (Harris-Perry 2011; Brown 2014; Lemi and Brown 2019). Schneider and Bos (2011) find that Black politicians constitute a psychologically distinct subtype, with their own stereotype content separate from general stereotypes of Black Americans; a finding that carries additional complexity when the candidate is also a woman. The result is that Black women candidates face a compounded, intersectionally structured set of voter expectations that shapes not only how they are evaluated but also which behaviors and rhetorical choices are legible as appropriate or authentic.¹

This pressure to perform manifests in two simultaneous normative regimes. First, the masculine-coded norms that structure political authority broadly reward toughness, assertiveness, and command. Second, the raced norms of respectability that govern which women are allowed entry into authority structures (Harris-Perry 2011). The double bind this creates for Black women candidates — between performing femininity and risking perceived lack of authority, or performing masculinity, and assertiveness at the risk of

¹This dynamic connects to Fenno's (2003) account of how Black politicians perform racial authenticity for both Black and White constituents in their district homestyle — a performance that complicates moderation, since candidates must neither perform race to an extreme degree that alienates cross-racial coalitions nor abandon group-linked credibility that maintains the base's support. Gillespie (2012) theorizes this tension directly in *The New Black Politician*, arguing that post-civil-rights Black candidates navigate a cross-pressured authenticity dilemma that has no parallel for White candidates.

perceived inauthenticity — is an evaluative constraint with no straightforward strategic escape, and one that White women candidates do not face in the same configuration.

Recent work has shown that this intersectional disadvantage is not uniform across all women of color. Snipes (2025) demonstrates that electoral patterns advantaging Democratic women in competitive districts are driven primarily by White women, with women of color — and particularly Black women — not benefiting from the same structural dynamics. Bejarano (2013) reinforces Crenshaw’s (1989) claim that the co-existence of racial and gender minority status produces interactive rather than additive effects that vary across subgroups, finding that these dynamics do not operate identically for Black women and Latinas. Chen and Sorensen (2025) identify a specific paradox facing Black women in Democratic primaries where voters perceive them as more liberal, which would ordinarily be an asset among Democratic primary electorates, but simultaneously rate them as less electable, undermining the electoral benefit of that ideological perception. The concept of electability, they argue, is raced and gendered in ways that effectively penalize Black women for the same ideological profile that advantages other candidates. Chirco and Buchanan (2022) add an aesthetic layer, showing that skin tone operates as an independent visual heuristic, with darker-skinned candidates perceived as less trustworthy and less competent by White voters even when controlling for other candidate characteristics.

2.3 Black Women’s Unique Strategic Environment

The stereotype landscape documented above does not simply act on Black women candidates from the outside; it shapes the strategic terrain they must actively navigate from within. Hancock (2004) traces how the “welfare queen” archetype, constructed through decades of racialized political discourse, ascribed a politics of disgust specifically to Black women’s relationship with the state. This culturally embedded image does not disappear when Black women run for office. Black women candidates are routinely assumed by voters to prioritize issues of racial group advocacy, poverty, and social welfare over economic management and

foreign policy, regardless of their actual legislative records or stated platforms (McDermott 1998; Carew 2016; Sigelman et al. 1995).

Gay and Tate (1998) demonstrate through national survey data that racial and gender identities among Black women are mutually reinforcing rather than hierarchically ordered, meaning both remain simultaneously salient in ways that make selective identity suppression unavailable as a strategy — in contrast to Latina candidates who can tap into “representational flexibility” (Bejarano 2013). Orey and Zhang (2019) find that among Black millennial voters, darker-skinned Black women candidates are perceived as holding more progressive and Afrocentric policy positions than their lighter-skinned counterparts — a finding that extends to a comparative group context, showing that colorism is distinctly gendered. Black women experience skin-tone discrimination more often than Black men, and the intersection of gender and skin tone shapes the evaluative standards applied to Black women candidates in ways that have no parallel for other candidate groups (Lemi and Brown 2021).

Black women who seek elected office do so with full awareness of this environment. Dowe (2020) theorizes what she calls an “ambition of marginality” — a form of political agency built in deliberate response to structural and stereotypical barriers. This awareness is consequential for campaign strategy of Black women candidates. They enter electoral contests already calculating how to manage perceptions they did not create. Harris-Perry’s (2011) concept of the “crooked room” describes this phenomenon as the distorted political and cultural environment Black women must navigate, requiring them to straighten themselves to be seen as legitimate political actors in a system whose evaluative standards work against them inherently. Santia, Bauer and Gonzalez (2025) complement this picture by showing that women of color navigate their campaign communications strategically, relying more heavily on positive emotional appeals than negative ones — evidence that the rhetorical choices of Black women candidates are already shaped by awareness of the distinctive evaluative standards they face.

The closest analog in the existing literature is the deracialization strategies documented

among Black candidates. Deracialization — the deliberate de-emphasis of race-specific appeals in favor of universalist, cross-racial policy language — represents one established way in which Black candidates have responded to the electoral pressures of running in majority-White or racially mixed constituencies (McCormick and Jones 1993; Gillespie 2012). The logic runs parallel to what ideological moderation theory predicts for the primary-to-general transition; candidates soften the most electorally threatening dimension of their identity in order to broaden their coalition.

3 Theory

Previous scholarship establishes that Black women candidates face an environment with ideological stereotypes, electability discounts, and phenotypic evaluations that combine to create unique pressures for moderation. Building on this foundation, I advance three predictions about how this pressure manifests in Black women's campaign rhetoric across the primary-to-general election transition.

The first prediction follows directly from the baseline logic of moderation. Spatial models of two-stage electoral competition demonstrate that candidates face cross-pressured incentives — appealing to their partisan primary base while simultaneously positioning for the broader general electorate (Adams and Merrill 2003, 2014). Adams and Merrill (2003) show that voters' nonpolicy characteristics, including race and partisanship, shape candidate positioning strategies even when candidates cannot directly manipulate those characteristics — a finding with direct implications for candidates who, like Black women congressional candidates, face electorates whose nonpolicy evaluations of them are shaped by entrenched racial and gender stereotypes. The basic assumption at play is that moving from ideologically mobilized primary appeals toward more centrist general election messaging will maximize support across both stages of the election. This moderation incentive is at minimum as strong for Black women candidates as for other candidates. I therefore predict:

H1 — Ideological Moderation: *Black women congressional candidates will moderate their campaign rhetoric from the primary to the general election.*

The second prediction extends this logic. If the stereotypical burden Black women carry is more intense than that of other candidate groups, then the magnitude of their post-primary moderation should exceed that of candidates who carry only one dimension of that burden, or neither. Black men face racial stereotyping but not the gendered electability discount. White women face gender stereotyping but not the racial ideological attribution. Black women face both simultaneously, and to a greater degree than Latina candidates (Bejarano 2013). Taking these assumptions, I predict:

H2 — Comparative Moderation Hypothesis: *The extent of rhetorical moderation among Black women congressional candidates will be greater than that of Black men, White women, White men, Latina women, and Latino men after the primary*

The third prediction concerns not the intensity but the content of rhetorical change. The ideology-electability paradox documented by Chen and Sorensen (2025) operates through a specific mechanism: Democratic primary voters perceive Black women as more liberal and simultaneously less electable, meaning the identity-linked appeals that may help Black women mobilize their primary base become liabilities in the general election. This creates an incentive that goes beyond ideological moderation to a form of deracialization. Because identity-linked appeals from Black women candidates risk reinforcing voter stereotypes, I predict that:

H3 — Issue Shift Hypothesis: *Black women congressional candidates will reduce their use of identity-linked appeals after the primary to a greater degree than candidates in other identity groups.*

H2 and H3 are comparative by design, following relevant scholarship in race and ethnic politics that seek to isolate what is distinctive about Black women's political environment through cross-group analysis rather than description alone (Carew 2016; Lemi and Brown

2021; Silva and Skulley 2019; Shah, Scott and Juenke 2019; Holman and Schneider 2018).

4 Data

This study draws on an original panel dataset combining candidate-level demographic and electoral characteristics with time-stamped campaign tweets from the 2024 U.S. House election cycle. The dataset includes over 17,000 tweets from the campaign accounts of 128 Democratic House candidates, collected from January 1, 2024 through Election Day, November 5, 2024.

4.1 Candidate Metadata

I constructed a master candidates-list of Democratic U.S. House candidates who ran and won their primary contests in 2024. The analysis is limited to Democratic candidates to maintain the partisan context constant and reduce heterogeneity driven by ideological differences between parties – a decision supported by evidence that moderation patterns differ between parties and are more strongly documented among Republicans who have become increasingly extreme in their primaries (Wittenbrink 2022). Candidates who did not maintain an active campaign Twitter account or who posted no tweets during the election cycle were excluded from the sample.

Table 1: Distribution of Tweets Across Groups

Race-Gender Group	Number of Tweets
White Men (WM)	6294
Black Men (BM)	3203
Black Women (BW)	2930
Latina Women (LW)	2419
White Women (WW)	1715
Latino Men (LM)	976

Candidate identifiers, district information, and campaign account handles were compiled by hand and verified against campaign websites and account bios. Race and gender identity

classifications were coded using publicly available congressional caucus membership records and candidate biographies. Candidates were categorized into six identity groups: Black women (BW), Black men (BM), White women (WW), White men (WM), Latina women (LW), and Latino men (LM). These classifications form the identity grouping variable used throughout the analysis.

Incumbency status and primary election dates were drawn from Ballotpedia and validated against official congressional listings on House.gov. Candidates are coded as incumbents if they held the seat at the time of the election and were seeking reelection. District partisanship is measured using the Cook Partisan Voting Index (Cook PVI), which captures each district’s baseline partisan lean relative to the national vote (Wasserman 2025). Cook PVI is included as a control variable to account for the structural electoral context in which candidates operate.

Table 2: Candidate and Tweet Descriptive Statistics by Identity Group

Identity Group	Mean Cook PVI	Mean Tweet Vol.	Mean Ideology	SD Ideology
Black Women	-22.8	244.2	-0.001	0.326
Black Men	-13.5	139.3	-0.066	0.387
Latina Women	-10.7	201.6	-0.062	0.434
Latino Men	-13.7	81.3	-0.115	0.372
White Women	-9.9	114.3	-0.057	0.405
White Men	-10.8	116.6	-0.006	0.378

4.2 Candidate Text Data

Tweets were collected programmatically from candidates’ official campaign accounts using a third-party Twitter API provider.² Only campaign accounts are included; official government accounts are excluded because of Hatch Act prohibitions on using official channels for

²Tweet data were retrieved using RapidAPI’s Twitter endpoint (<https://rapidapi.com>), which provides API-equivalent access to Twitter/X data. Only original tweets authored by the candidate account were retained; retweets were excluded to avoid incorporating third-party messaging, while quote-tweets were kept because they contain original candidate-authored commentary.

campaign communication.³ In cases where incumbents tweeted exclusively from their government accounts, those candidates are omitted from the sample.

Each tweet was timestamped and aligned to the candidate’s primary election date, producing a running variable *DaysFromPrimary* that measures each tweet’s distance in days from the primary date cutoff. This variable is used in both the DiD and RD analyses described below. For panel models, tweets are aggregated to the candidate-week level to prevent high-volume accounts from exerting disproportionate influence on estimates.⁴

In addition to ideological scores, the dataset includes a measure of identity-linked appeals constructed using a theory-driven dictionary of race- and gender-related terms. This measure captures the frequency of identity-framed rhetoric in each tweet and is used to test H3.

5 Research Design

5.1 Ideology: Conceptual Definition and Measurement

In this paper, I refer to ideology as the expressed liberal-conservative orientation of candidate campaign rhetoric. Specifically, ideology is the degree to which a candidate’s public communication emphasizes progressive or conservative policy positions, frames issues in ways associated with the left or right, and signals ideological alignment through language choices. This framing is distinct from legislative definitions of ideology inferred from roll-call voting behavior and stated policy positions through press releases, campaign websites, or prepared speeches — communication that passes through professionalized production pathways and is more susceptible to strategic curation before reaching the public. Twitter data, by contrast, are comparatively raw: frequent, reactive to real-time events, and produced

³The Hatch Act of 1939 (5 U.S.C. §§ 7321–7326) restricts federal employees and officeholders from engaging in partisan political activity using government resources, including official social media accounts.

⁴Tweet volume varied substantially across candidates. Jasmine Crockett (TX-30) accounted for 1,119 of 17,537 total tweets in the dataset — approximately 6.4% of all observations — making her the highest-volume candidate by a substantial margin. Within the Black women subsample specifically, Crockett accounted for 1,119 of 2,930 tweets (38.2%). Week-level aggregation mitigates but does not fully eliminate the influence of high-volume accounts, and results should be interpreted with this imbalance in mind.

at a volume and pace that limits the degree of deliberate ideological filtering. This makes Twitter a useful observational window for detecting genuine rhetorical shifts.

To score the ideological content of each tweet, I apply zero-shot classification using the BART-large-MNLI transformer model (Lewis et al. 2020), implemented via the Hugging Face transformers library (Wolf et al. 2020). Following Bailey (2023) and Bailey and Reese (2025), I treat candidate-authored text as the preferred basis for ideology measurement for the following reasons: 1) candidates control their language directly, 2) that language bears a close relationship to how a broad range of voters perceive their ideological positioning, and 3) text-based measurement is not constrained by the roll-call agenda set by party leaders. Two notable advantages of this approach are the ability to measure ideology for non-incumbent challengers who have no voting record, and to capture ideological positioning on issues that may rarely or never reach the floor for a vote.

I chose a zero-shot classification over supervised ideology classifiers for two reasons specific to this study's design. First, existing supervised classifiers trained on labeled political text from prior election cycles risk domain mismatch when applied to 2024 congressional campaign tweets, which differ in register, topic, and rhetorical context from prior corpora. Zero-shot classification applies a consistent scoring procedure across all candidates and time periods without inheriting prior-cycle calibration assumptions. Second, constructing a supervised classifier for the 2024 context would have required either an existing labeled dataset of 2024 congressional campaign tweets (which does not currently exist in the published literature) or an original hand-coded corpus followed by the computational resources necessary to fine-tune a large language model on that labeled data. Both represent resource demands that exceed the scope of this project. Given these considerations, a zero-shot classification via BART-large-MNLI is the best available alternative as it leverages natural language inference without requiring domain-specific training data and has been increasingly adopted for political text analysis (Burnham et al. 2025; Heseltine and von Hohenberg 2024).

The probability scores output by the model are converted to a continuous ideology score by weighting each label: liberal receives -1 , moderate 0 , and conservative $+1$. The weighted sum produces a score ranging from -1 (strongly liberal) to $+1$ (strongly conservative), with values near zero indicating ideologically neutral or ambiguous content. For example, a tweet scored 80% liberal, 15% moderate, and 5% conservative yields a score of -0.75 .

Table 3: Illustrative Tweet Ideology Scores

Author	Tweet	Liberal ($\times -1$)	Moderate ($\times 0$)	Conservative ($\times 1$)	Ideology Score
Angie Craig	<i>“The endgame for Republicans has always been the same: a nationwide abortion ban. No matter what they call it, their legislation would take away the freedom of every family, in every state, to make their own decisions at the doctor’s office.”</i>	0.955	0.040	0.005	-0.950
Valerie Foushee	<i>“It’s not happening lol. I know my limits. This is one of them.”</i>	0.060	0.875	0.065	-0.005
Valerie Foushee	<i>“States run our elections. They all have their own rules of how you get on the ballot. It is the reason that third party candidates aren’t on ballots in all 50 states.”</i>	0.175	0.475	0.350	$+0.175$

To illustrate what these scores capture in practice, Table 3 presents three examples drawn from the dataset. A strongly liberal tweet (score: -0.91) from Angie Craig reads: “The endgame for Republicans has always been the same: a nationwide abortion ban. No matter what they call it, their legislation would take away the freedom of every family, in every state, to make their own decisions at the doctor’s office.” A near-neutral tweet (score: -0.01) from Valerie Foushee reads: “It’s not happening lol. I know my limits. This is one of them.” A conservative-coded tweet (score: $+0.44$) also from Foushee reads: “States run our elections. They all have their own rules of how you get on the ballot. It is the reason that third party candidates aren’t on ballots in all 50 states.” The third example illustrates a limitation of the measure — tweets that are politically liberal in context but expressed through procedural or institutionally neutral language may receive conservative-leaning scores. This limitation is further addressed in the validation analysis.

5.2 Measurement Validation

5.2.1 Inter-Rater Reliability

To assess the validity of the BART-MNLI ideological scoring measure, I conducted a hand-coded inter-rater reliability validation on a random sample of 100 tweets drawn from the full tweet corpus. For the hand-coded scoring, I rated each of the sample tweets categorically as either liberal, conservative, or neutral. Since the ideology measure is a continuous score on a liberal-to-conservative continuum, the validation would consist of measuring the agreement between the human categorization and the continuous score, with negative values indicating liberal rhetoric and positive values indicating conservative rhetoric.

To compare the two measures, I converted the continuous BART-MNLI scores into the same three categories using neutral-band thresholds ($|\text{score}| > 0.05$ through 0.30) and selected the threshold that maximized Cohen's κ , following standard measures of inter-rater agreement (Landis and Koch 1977). The optimal threshold of $|\text{score}| > 0.30$ yielded an overall percent agreement of 80.0% and $\kappa = 0.546$, indicating moderate agreement between the automated and human rating. Among the subset of tweets that both the human coder and the model independently classified as ideological rather than neutral, directional agreement reaches 95.0%, suggesting that BART-MNLI reliably recovers the *direction* of ideological content when it detects ideology at all.

The model's primary limitation is a tendency to classify liberal tweets as neutral: 41.4% of hand-coded liberal tweets received a neutral score. These cases are consistent with liberal candidates' documented use of constituency service language, endorsements, and issue framing that lacks the explicit partisan markers BART-MNLI uses to detect ideology. Mean BART-MNLI scores are substantively ordered in the expected direction across categories (liberal: -0.397 ; neutral: -0.014 ; conservative: $+0.376$), and a two-sample t -test confirms that scores for hand-coded liberal tweets are significantly more negative than those for hand-coded neutral tweets ($t = -6.165$, $p < 0.001$). These results provide reasonable

empirical support for the validity of BART-MNLI as an automated ideology measure, while acknowledging its tendency to underdetect liberal rhetoric when candidates use implicit rather than explicitly partisan language.

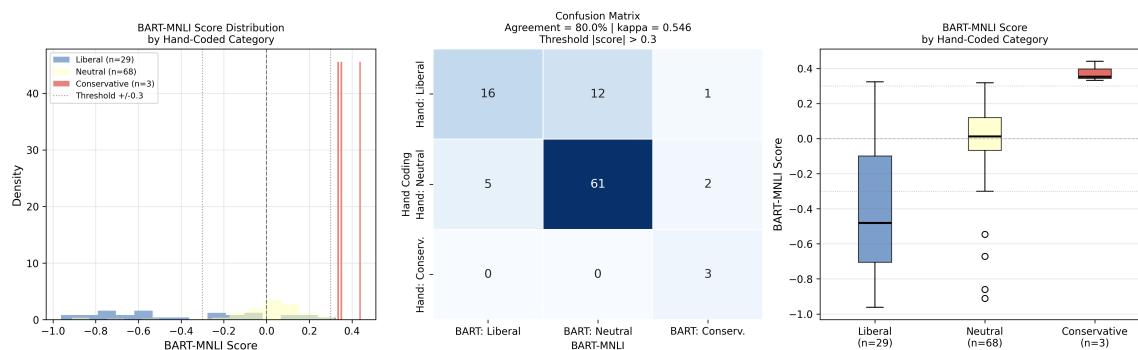


Figure 1: BART-MNLI validation results. *Left*: Score distribution by hand-coded category. *Center*: Confusion matrix at threshold $|\text{score}| > 0.30$ (agreement = 80.0%, $\kappa = 0.546$). *Right*: Box plots of BART-MNLI scores by hand-coded category.

5.2.2 Validity with Legislative Ideology Scores

As a second validation check, I assess whether the BART-MNLI measure exhibits convergent validity with established legislative ideology scores. Specifically, I correlate each candidate's BART-MNLI baseline ideology — defined as their mean *IdeologyScore* averaged across all tweets in the corpus — with their DW-NOMINATE first-dimension score and their Nokken-Poole first-dimension score, two widely used measures of legislator ideal points derived from roll-call voting records (Poole and Rosenthal 1985; Nokken and Poole 2004). DW-NOMINATE and Nokken-Poole scores are scaled from -1 (most liberal) to $+1$ (most conservative), matching the directional logic of the BART-MNLI measure. A positive correlation is theoretically expected if both instruments are capturing the same underlying ideological dimension.

I merge the BART-MNLI baselines for the candidates in the corpus with 118th and 119th Congress DW-NOMINATE records using candidate name. Of the 128 candidates, 117 are successfully matched, with the 11 unmatched cases being challengers with no floor voting record, and therefore no scorable DW-NOMINATE estimate. Among the candidates, the

BART-MNLI baseline correlates positively with both DW-NOMINATE ($r_{\text{Pearson}} = 0.236$, $p = 0.012$; $r_{\text{Spearman}} = 0.265$, $p = 0.004$) and Nokken-Poole scores ($r_{\text{Pearson}} = 0.258$, $p = 0.006$; $r_{\text{Spearman}} = 0.317$, $p < 0.001$), in the expected direction. Both the hand-coded inter-rater reliability results and the convergent validity analysis with legislative ideology scores provide reasonable empirical support for treating the BART-MNLI ideology measure as a valid instrument for detecting ideological direction in candidate social media rhetoric.

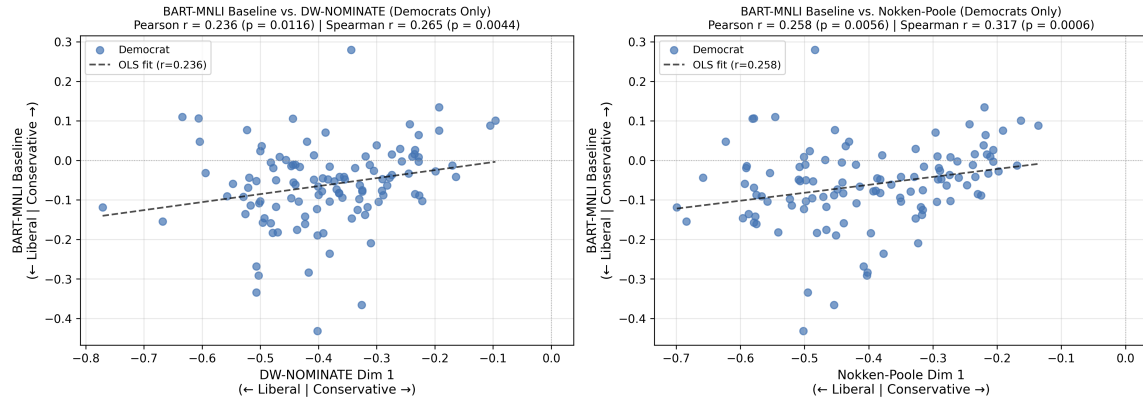


Figure 2: BART-MNLI baseline ideology correlated with DW-NOMINATE (left) and Nokken-Poole (right) first-dimension scores, Democrats only ($N = 114$). Dashed line is OLS fit; both axes are oriented liberal (left) to conservative (right).

5.3 Analytical Strategy

5.3.1 Difference-in-Differences

The primary analytical design is a two-way fixed-effects Difference-in-Differences model estimated at the tweet level and the candidate-week level. The primary specification is:

$$\text{IdeologyScore}_{it} = \beta_1 (\text{BW}_i \times \text{PostPrimary}_{it}) + \alpha_i + \gamma_t + \varepsilon_{it} \quad (1)$$

where α_i are candidate fixed effects, γ_t are relative-week fixed effects, and standard errors are clustered at the candidate level. The coefficient of interest, β_1 , captures whether Black women's ideology scores shift differentially after the primary relative to other candidates. A negative and statistically significant β_1 would support H1 (moderation) and H2 (comparative

moderation). I also estimate a group comparison model interacting *PostPrimary* with all six identity indicators, using Black women as the reference category, to directly assess H2.

To ensure results are not sensitive to a single operationalization, I estimate three parallel models differing only in the dependent variable. The first uses the raw continuous ideology score, capturing directional rhetorical positioning. The second uses the absolute distance from the ideological center, $|\text{IdeologyScore}_{it}|$, measuring moderation as movement toward zero regardless of direction. The third uses a conservative-intensity score that captures only positive ideological signals, testing whether post-primary moderation occurs specifically through increased conservative-aligned framing. All three specifications use identical treatment timing, group interactions, and fixed-effects structures.

5.3.2 Regression Discontinuity

As a complementary design, I estimate a local polynomial Regression Discontinuity using the primary date as the cutoff and *DaysFromPrimary* as the running variable, implemented via the `rdrobust` package in R (Calonico, Cattaneo and Titiunik 2014). Rather than comparing the full pre- and post-primary periods, the RDD asks whether there is a sharp, discontinuous shift in ideological tone precisely at the primary cutoff. The bandwidth is set to 90 days on each side of the primary. The dependent variable is the raw continuous ideology score. Results are reported alongside the DiD estimates as a robustness check on the main findings.

One potential threat to the DiD design is anticipatory adjustment (Lee and Lemieux 2010), where, in this context, a candidate who expects to win their primary may begin moderating before the primary date and compress the measured post-primary effect toward zero. To assess this, I plot pre-primary ideological trajectories separately for incumbents and non-incumbents — the candidates most and least likely to anticipate a primary win, respectively (Freyaldenhoven, Hansen and Shapiro 2019). If anticipatory adjustment were driving the estimates, we would expect incumbents to show steeper pre-primary slopes than

non-incumbents. Figure 3 shows no systematic divergence between the two groups across any identity group, suggesting anticipatory adjustment does not materially bias the DiD estimates.

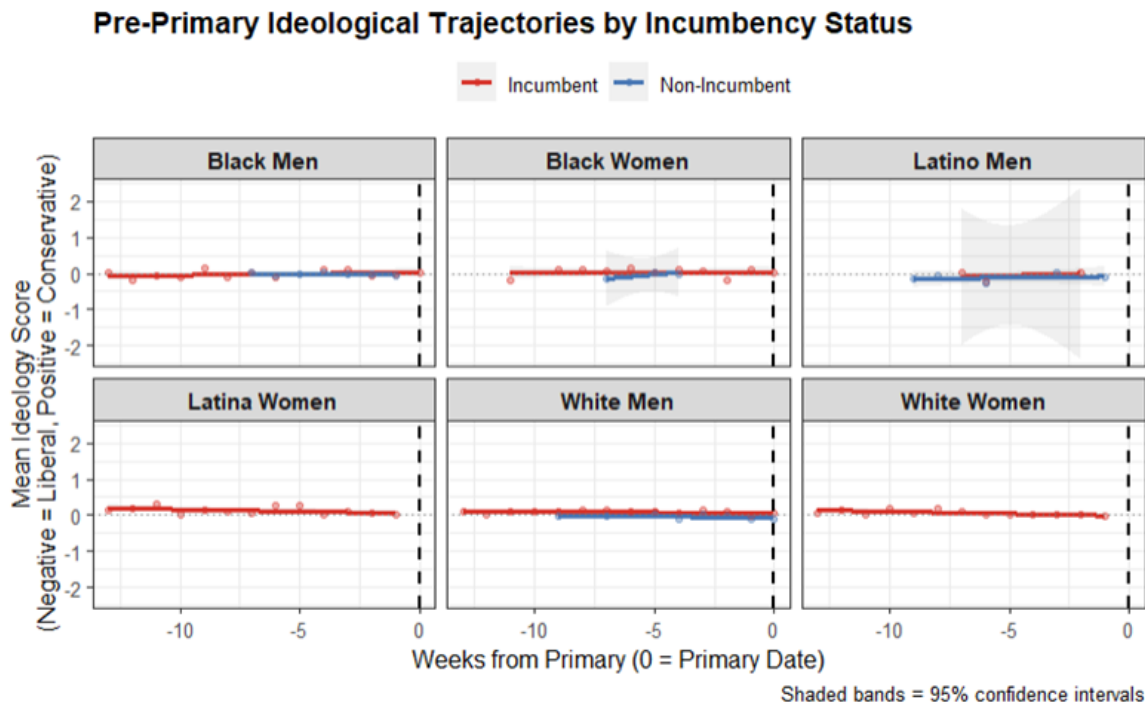


Figure 3: Pre-Primary Ideological Trajectories by Incumbency Status. Each panel displays mean weekly ideology scores in the pre-primary period for incumbents (red) and non-incumbents (blue) within each identity group. The dashed vertical line marks the primary date (week 0). The y-axis = the mean BART-MNLI ideology score, where negative values indicate liberal framing and positive values indicate conservative framing.

5.3.3 H3: Identity Appeals Analysis

H3 is tested using a DiD framework applied to identity-linked rhetoric. The dependent variable is constructed from a theory-driven dictionary of identity and representation terms capturing racial and ethnic group references (“Black community,” “women of color,” “immigrant families”), justice frames (“racial equity,” “voting rights,” “criminal justice reform”), representation language (“fight for,” “historic first,” “lived experience”), intersectional signifiers (“marginalized,” “women of color”), and empowerment and solidarity frames (“uplift,” “amplify,” “coalition”). Approximately 29.8% of tweets in the full sample contain at least

one identity-linked term, confirming that identity appeals are a meaningful and measurable feature of candidate rhetoric in this corpus. The primary specification mirrors Equation (1) with *Identity_flag* as the dependent variable.

6 Results

6.1 H1: Do Black Women Congressional Candidates Moderate After Primaries?

The evidence does not support the Ideological Moderation Hypothesis. Across three operationalizations of rhetorical ideology — directional score, absolute distance from center, and conservative intensity — Black women candidates show no statistically significant post-primary shift in either direction. Table 4 displays the within-candidate DiD estimate for the $BW \times PostPrimary$ interaction is $+0.027$ ($SE = 0.036$, $p > 0.10$), indicating a slight positive movement in ideological score after the primary, but one that is both substantively negligible and statistically indistinguishable from zero. The absolute distance model returns a coefficient of -0.001 ($SE = 0.013$), and the conservative intensity model returns $+0.010$ ($SE = 0.024$). None approach conventional thresholds of significance. The event-study plot confirms the absence of a systematic trend: ideology scores are distributed evenly across the pre- and post-primary window, with no visible break at the primary cutoff (Figure 4).

Table 5 shows the unadjusted RDD for Black women finds no significant shift at the primary cutoff ($RD = -0.024$, $p = 0.334$), consistent with the DiD results. When controls for incumbency and district partisanship are added, the estimate becomes statistically significant and reverses direction ($RD = -0.181$, $p = 0.023$), suggesting a liberal rather than conservative shift precisely at the cutoff. However, this counterintuitive result appears to be driven by the control variables rather than a genuine discontinuity in rhetoric. A stratified bootstrap equalizing each identity group to 12 candidates over 500 iterations yields an average $BW \times PostPrimary$ coefficient of 0.028, with no iteration reaching statistical significance. Neither the RDD nor the bootstrap provides evidence of post-primary

Table 4: DiD Estimates of Post-Primary Ideological Change for Black Women (H1)

	(1) Ideology Score	(2) Absolute Distance	(3) Conservative Intensity
BW \times PostPrimary	0.027 (0.036)	-0.001 (0.013)	0.010 (0.024)
PostPrimary \times Incumbent	-0.050 (0.027)		
PostPrimary \times Cook PVI	0.000 (0.001)		
<i>Fixed Effects</i>			
Candidate	Yes	Yes	Yes
Relative Week	Yes	Yes	Yes
<i>Note:</i> Standard errors in parentheses, clustered at the candidate level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. No coefficients approach conventional thresholds of significance.			

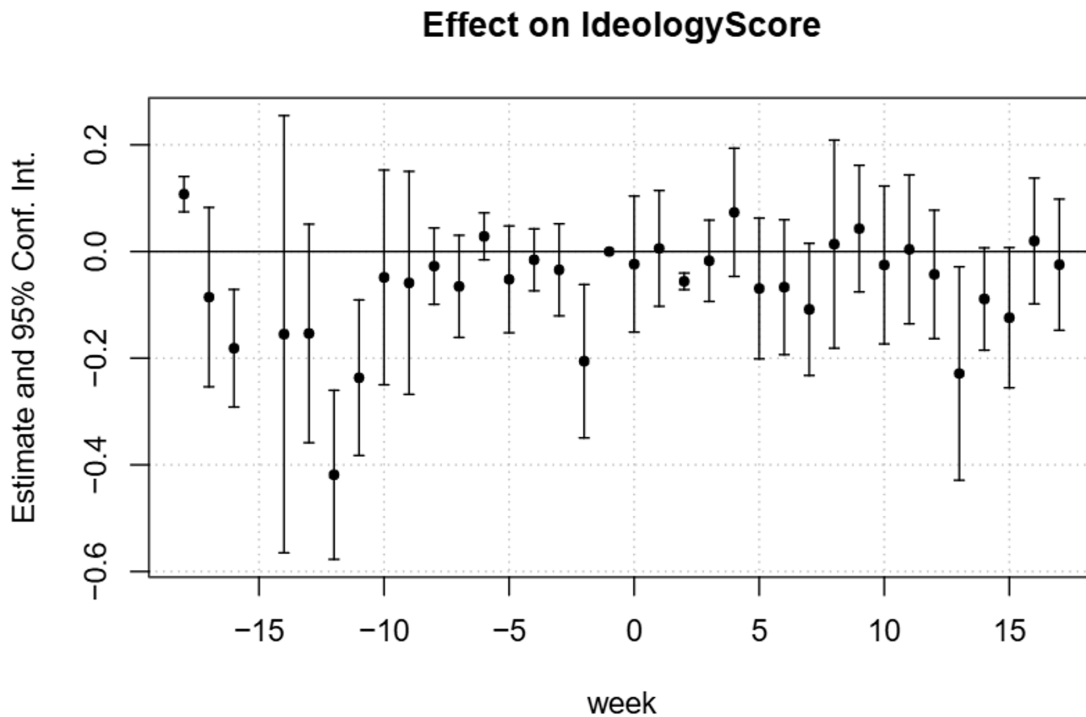


Figure 4: Event Study: Effect on *IdeologyScore*. Each point represents the week-specific DiD estimate for Black women relative to the comparison group, with bars indicating 95% confidence intervals. Week 0 denotes the primary date.

moderation. **H1 is not supported.**

Table 5: Regression Discontinuity Estimate at the Primary Cutoff: Black Women (H1)

	Estimate	z	$p > z $	95% C.I.
RD Effect	-0.181**	-2.270	0.023	[-0.414, -0.030]
<i>Design parameters</i>				
Obs. (left / right)	444 / 712			
Eff. obs. (left / right)	89 / 134			
Bandwidth h / b	14.075 / 28.024 days			
Kernel	Triangular			
Polynomial order	1			

Note: Covariate-adjusted sharp RD (rdrobust), MSE-optimal bandwidth, nearest-neighbor variance estimation. The negative estimate reflects a leftward — not rightward — shift at the primary cutoff, opposite to the moderation prediction. This result appears to be covariate-driven rather than a genuine discontinuity; the unadjusted estimate is -0.024 ($p = 0.334$). H1 is not supported.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.2 H2: Do Black Women Moderate More Than Other Groups?

The evidence does not support the Comparative Moderation Hypothesis. The $BW \times PostPrimary$ interaction in the comparative DiD — which tests whether Black women’s post-primary rhetorical shift differs from the pooled average of all other groups — is $+0.021$ ($SE = 0.040$, $p > 0.10$) at the tweet level and $+0.052$ ($SE = 0.037$, $p > 0.10$) in the candidate-week panel (Table 6). The fully-specified group model, which estimates each group’s post-primary shift relative to Black women as the reference category, finds that all five comparison groups show negative coefficients — meaning they moved slightly more in the liberal direction post-primary than Black women — but none of these differences reach significance. There is a marginal finding that Latina women show greater post-primary movement toward the ideological center than Black women on the absolute distance measure ($PostPrimary \times LW = -0.036$, $p < 0.05$), but this is opposite to the H2 prediction since Latina women moderated *more*, not less.

Descriptively, Black women show among the smallest mean absolute shifts across groups (0.107), compared to 0.136 for White women and 0.094 for Black men. The group trajectory

Table 6: DiD Estimate: Black Women’s Post-Primary Ideological Shift Relative to All Other Groups (H2)

	(1) Tweet-Level	(2) Candidate-Week
BW × Post-Primary	0.021 (0.040)	0.052 (0.037)
Candidate FE	Yes	Yes
Relative Week FE	Yes	Yes
Observations	8,516	1,701
R ²	0.079	0.251
Within R ²	0.000	0.001

Note: Standard errors in parentheses, clustered at the candidate level. Black Women (BW) is the treatment group; the coefficient captures whether Black Women’s post-primary ideological shift differs from the pooled average of all other identity groups. Column (1) estimated at the tweet level; Column (2) aggregates tweets to the candidate-week. Neither estimate approaches conventional thresholds of significance.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

plots confirm this pattern — no group exhibits a clear discontinuous shift at the primary cutoff, and Black women’s trajectory is broadly indistinguishable from the pooled trend across Democratic House candidates. The RDD estimates across all six groups are uniformly null, with point estimates ranging from -0.024 (Black women) to $+0.099$ (White women) and confidence intervals spanning zero in every case. **H2 is not supported.**

Table 7: Group Comparison DiD: Post-Primary Ideological Change Relative to Black Women (H2)

	(1) Ideology Score	(2) Ideology Score	(3) Conservative Intensity
<i>Reference group: Black Women</i>			
Post-Primary × Black Men	−0.024 (0.046)	0.033 (0.020)	0.005 (0.027)
Post-Primary × Latino Men	−0.025 (0.055)	0.024 (0.023)	−0.001 (0.028)
Post-Primary × Latina Women	−0.030 (0.045)	−0.036* (0.014)	−0.028 (0.027)
Post-Primary × White Men	−0.021 (0.041)	0.001 (0.016)	−0.010 (0.025)
Post-Primary × White Women	−0.008 (0.048)	−0.001 (0.022)	−0.005 (0.032)
Candidate FE	Yes	Yes	Yes
Relative Week FE	Yes	Yes	Yes
Observations	8,516	8,905	8,905
R ²	0.079	0.098	0.104
Within R ²	0.000	0.001	0.000

Note: Black Women is the reference category; coefficients for all other groups reflect their post-primary ideological shift *relative to* Black Women. A negative coefficient indicates the comparison group shifted more in the liberal direction post-primary than Black Women. Column (1) uses the raw continuous ideology score; Column (2) uses absolute distance from the ideological center ($|\text{Ideology Score}|$) as a measure of moderation; Column (3) uses conservative rhetorical intensity ($\max(\text{Ideology Score}, 0)$).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

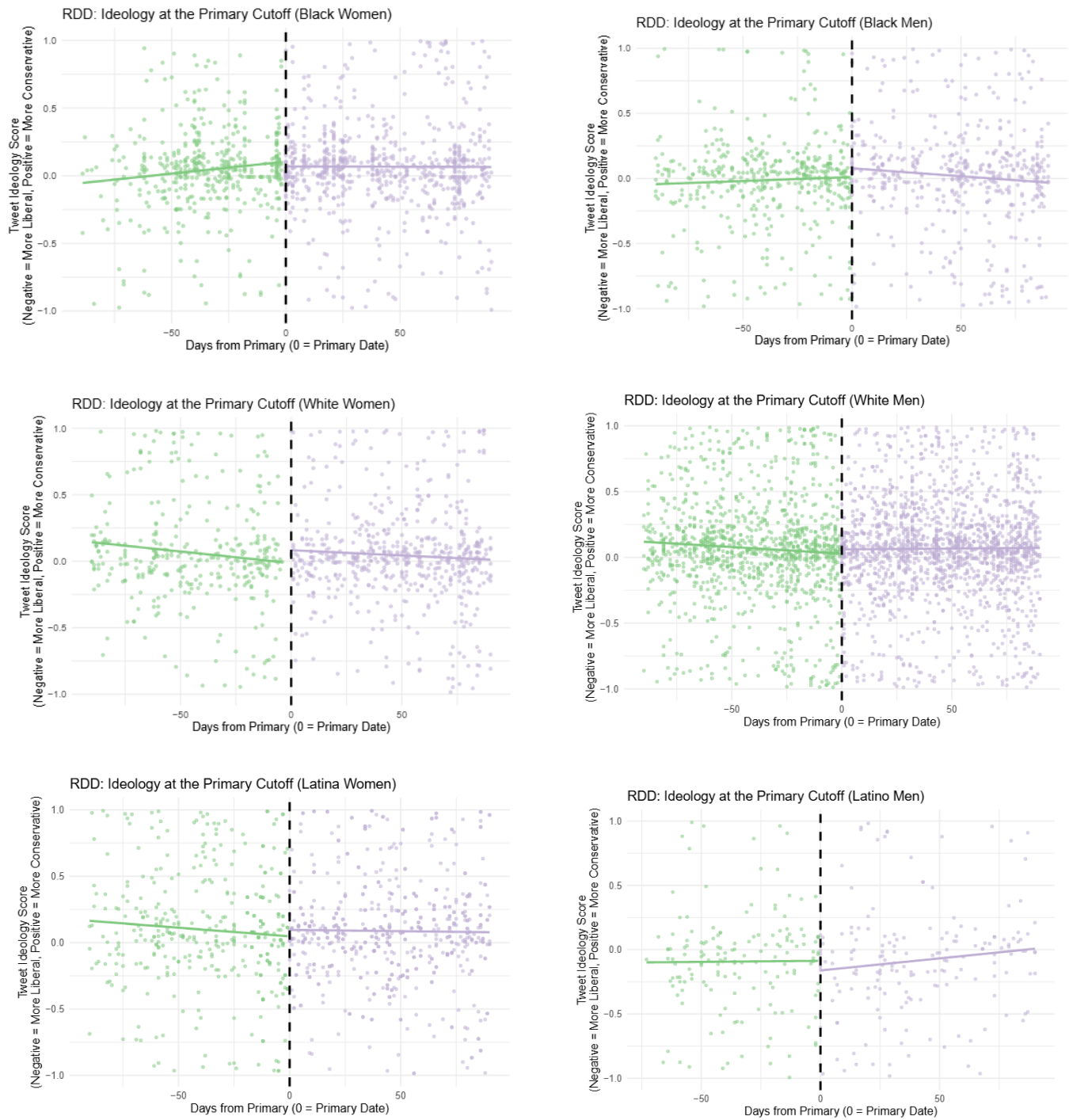


Figure 5: Regression Discontinuity Plots by Identity Group (H2). Each panel displays the local polynomial fit on either side of the primary election cutoff (day 0) for tweet-level ideology scores. No group exhibits a statistically significant discontinuous shift at the primary cutoff.

6.3 H3: Do Black Women Reduce Identity Appeals After Primaries?

Domain	Example terms
Racial/ethnic identity	<i>black, african american, latino, latina, aapi, native american, people of color, bipoc</i>
Gender identity	<i>woman, women, girls, mother, daughter, sister, women's rights, lgbtq, transgender, pride</i>
Justice & rights	<i>racial justice, racial equity, civil rights, voting rights, criminal justice reform, equal pay</i>
Representation	<i>first black, first woman, historic, trailblazer, fight for, lived experience, look like me</i>
Intersectional frames	<i>black woman, women of color, intersectionality, multiply marginalized, systemic, structural</i>
Empowerment & solidarity	<i>uplift, amplify, empower, coalition, movement, inclusion, diversity, solidarity, belonging</i>
Community appeals	<i>our community, my people, underserved, underrepresented, working families, left behind</i>
Identity-linked policy	<i>reproductive rights, daca, dreamer, voter suppression, affirmative action, hate crimes</i>
Cultural signifiers	<i>juneteenth, kwanzaa, black history month, hispanic heritage month, pride month, ancestors</i>

Figure 6: Identity Appeal Dictionary: Domains and Example Terms (H3). Terms are matched using whole-word regular expressions. The binary dependent variable (*identity_flag*) equals 1 if a tweet contains at least one term from any domain; 0 otherwise. Approximately 29.8% of tweets in the full sample are flagged.

The Issue Shift Hypothesis is not supported. The DiD estimate (Table 8) for $BW \times PostPrimary$ on the binary identity appeal indicator is -0.002 ($SE = 0.048$, $p > 0.10$), and the week-level model on the continuous identity rate returns $+0.073$ ($SE = 0.049$, $p > 0.10$). Neither is significant, and the latter coefficient runs counter to the predicted direction. The group comparison model finds no significant difference between Black women and any other group on post-primary identity appeal usage. Pre-to-post comparison plots show that Black women maintain relatively stable identity appeal rates across the primary transition, with overlapping confidence intervals between periods for all groups. Figure 7 illustrates the stability visually — the weekly identity appeal rate for Black women shows no decline at

the primary cutoff, and the trajectory is broadly indistinguishable from the pattern observed in other groups.

A closer inspection of the identity appeal trajectories reveals a notable spike in Black women’s identity appeal usage at approximately week 13 post-primary (identity rate = 62% that week, compared to 25–30% in surrounding weeks). Examination of the tweets driving this spike reveals that it falls in early- to mid-June 2024, a period coinciding with late primary contests for several Black women candidates and the intensification of Kamala Harris’s presidential campaign activity — which appears to have animated identity-linked rhetoric among this group (Figure 9). **H3 is not supported.**

Table 8: DiD Estimates: Black Women’s Post-Primary Identity Appeal Usage Relative to All Other Groups (H3)

	(1) Tweet-Level Identity Flag	(2) Candidate-Week Identity Rate
BW × Post-Primary	−0.002 (0.048)	0.073 (0.049)
Candidate FE	Yes	Yes
Relative Week FE	Yes	Yes
Observations	8,905	1,821
R ²	0.094	0.202
Within R ²	6.6×10^{-7}	0.001

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

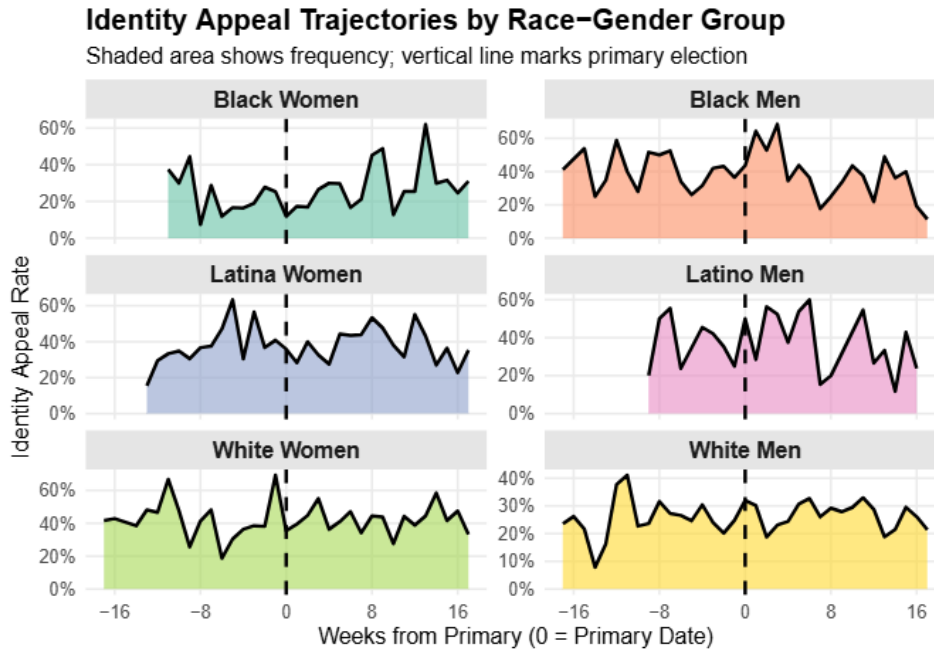


Figure 7: Identity Appeal Trajectories by identity Group (H3). Each panel displays the weekly identity appeal rate for each identity group across the primary election window. The shaded area shows appeal frequency; the dashed vertical line marks the primary election date (week 0).

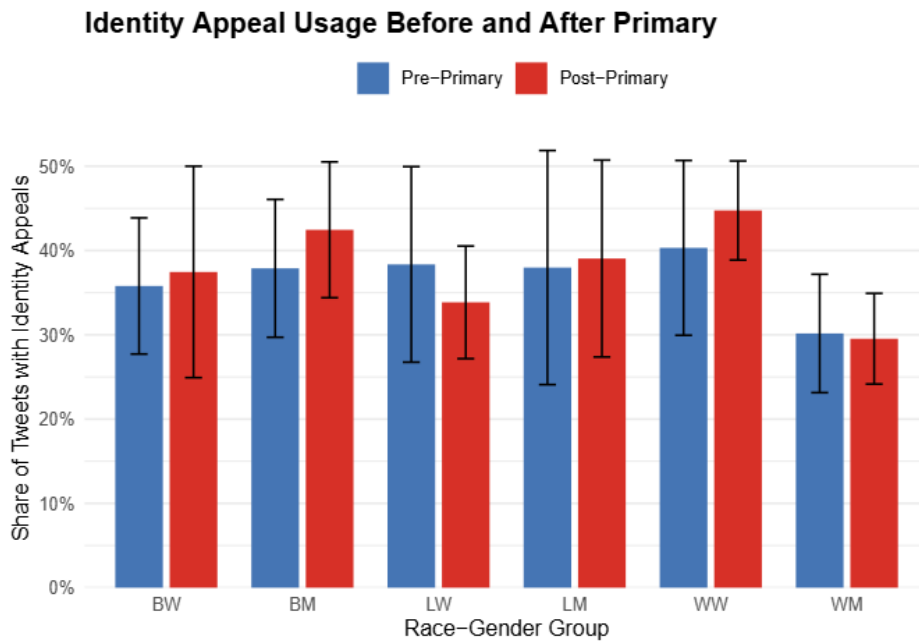


Figure 8: Identity Appeal Usage Before and After Primary by identity Group (H3).

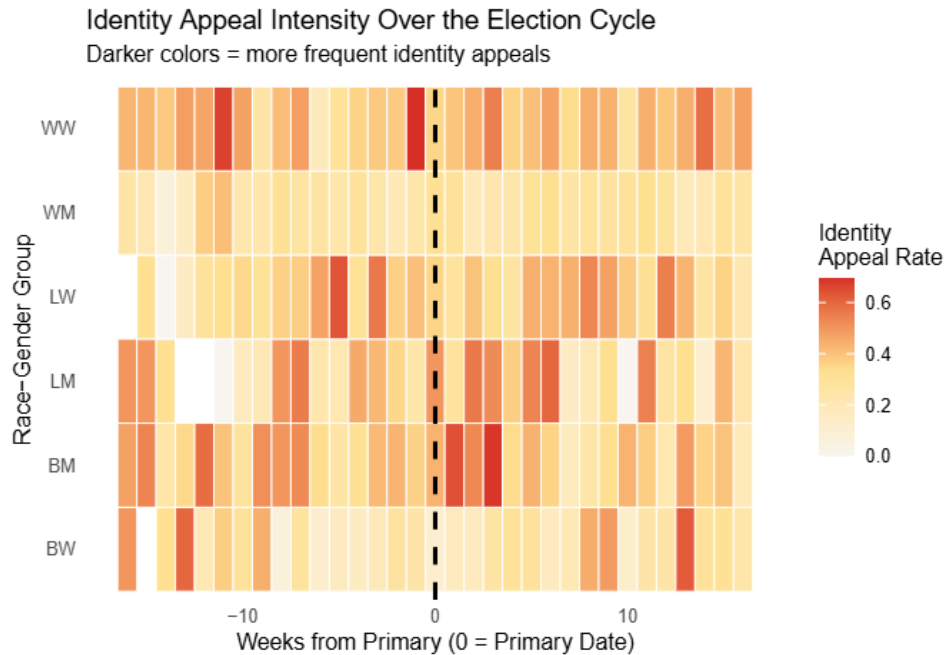


Figure 9: Identity Appeal Intensity Over the Election Cycle by identity Group (H3). Each cell represents the mean weekly identity appeal rate for a given group; darker shading indicates higher frequency of identity-linked rhetoric. The elevated intensity visible among Black Women (BW) and Black Men (BM) in the immediate post-primary weeks coincides with the onset of Kamala Harris’s presidential campaign in June 2024, which appears to have animated identity-linked rhetoric among Black candidates regardless of their own primary timing.

7 Conclusion

Black women are running for congressional office at historically unprecedented rates, yet the electoral environment they navigate remains structured by stereotype, skepticism, and a double bind that renders their ideological profile simultaneously an asset and a liability. In this paper I ask whether that burden produces a measurable behavioral response: do Black women candidates moderate their campaign rhetoric more sharply after winning primaries than candidates who do not carry this intersectional weight? The answer this study produces is clear. Across three hypotheses, two analytical strategies, and a range of operationalizations of ideology and identity, Black women congressional candidates in the 2024 election cycle did not systematically moderate their ideological rhetoric after winning

primaries, did not do so to a greater degree than comparison groups, and did not reduce their use of identity-linked appeals in the post-primary period.

The absence of differential moderation among Black women candidates could, in fact, be a substantive claim about how these candidates navigate electoral environments. The standard convergence model assumes that post-primary moderation is incentivized by the need to attract cross-partisan voters in competitive general elections. But Black women in this sample run overwhelmingly in heavily Democratic districts with constituencies where the general electorate is itself quite liberal and where the reputational cost of perceived ideological repositioning may outweigh the marginal electoral benefit of moving toward the center. Future work with a more diverse sample of Black women candidates in genuinely competitive districts may expand on this inquiry.

Black women's rhetoric is also more stable around the primary cutoff than several comparison groups, with Black men, White women, and Latina women all exhibiting larger mean absolute shifts. This stability may reflect several things simultaneously — that Black women have already negotiated the moderation-authenticity tradeoff through the candidate emergence process itself (Dowe 2020; Harris-Perry 2011); that Twitter-based rhetoric in safe Democratic districts is primarily constituent-service oriented rather than ideologically calibrated; or that Black women maintain their rhetorical positioning precisely because the electoral environment in which they disproportionately run does not reward deviation from it.

8 Limitations

Two limitations of this study warrant direct attention. First, the sample is structurally constrained. With twelve Black women candidates, all in safe Democratic districts, the analysis has limited statistical power to detect comparative moderation effects even if they exist, and the results cannot be generalized to Black women running in more competitive or ideologically heterogeneous environments without expanding the scope to prior election

years. Second, the ideology measure carries known limitations. The BART-large-MNLI zero-shot classifier, while consistently applied and partially validated, systematically underestimates the liberal intensity of tweets that are politically progressive in context but expressed through institutionally neutral language — a potential bias that may have worked against finding a result. If liberal rhetoric from Black women candidates was undercounted by the model, post-primary shifts in the liberal direction would be harder to detect.

9 Future Directions

This paper is best understood as a first step in a research program rather than a definitive test of a settled theory. The scope of the current study — twelve Black women candidates, one electoral cycle, a single text platform — reflects understandable constraints rather than the full ambition of the question. Expanding the corpus in future iterations would substantially improve the project’s inferential leverage. Incorporating campaign websites, press releases, floor speeches, and debate transcripts alongside Twitter data would produce a richer and more representative picture of candidate rhetoric across communication contexts that vary in audience, register, and strategic intent. A more expansive candidate pool extending to those who won primaries but lost general elections could better inform the model of whether these strategies are widespread, and whether they may show varying levels of success across race. The analysis could also be broadened to include state legislative races where many more Black women run for office and currently account for 5.4% of all state legislators (Center for American Women and Politics 2025).

Another meaningful contribution could be a supervised transformer model fine-tuned on a bespoke labeled dataset of political text drawn from the 2024–2026 political environment. This tailored, scalable method of scoring would be better positioned to detect the nuanced ideological signals this study attempts to measure, capturing the rhetorical conventions, policy debates, and identity-linked frames specific to this political moment in ways that a general-purpose zero-shot classifier cannot.

Beyond measurement, the theoretical architecture of the project could be strengthened before the empirical tests are run. As it stands, the study moves quickly from documented stereotypes to predictions about rhetorical behavior — something that may be putting the cart before the horse. A more complete version of this project would first establish that voters do in fact perceive Black women candidates as more ideologically extreme than their actual stances warrant, and would further parse whether that perception applies uniformly or varies by candidate characteristics. Not all candidates presumed to be more liberal than they are face the same electoral consequences — district competitiveness, the racial composition of primary and general electorates, institutional and party barriers to candidate entry, and the specific visibility of individual candidates all condition the degree to which ideological stereotyping translates into behavioral moderation pressure.

This connects to a dimension of the question this study cannot address quantitatively but which Brown and Lemi (2021) make impossible to ignore: the heterogeneity within the category of Black women candidates. Skin tone, natural versus straightened hair, professional presentation, speech patterns, perceived viability, and homestyle all shape how candidates are evaluated. The moderation incentive this paper theorizes is not uniform across Black women candidates; it is conditioned by the phenotypic, stylistic, and contextual signals that different candidates send to different audiences in different electoral environments. Capturing that variation requires not only larger data and better measurement, but qualitative work that can attend to the texture of how individual Black women candidates navigate the crooked room.

References

- Acree, Brice D. L., Justin H. Gross, Noah A. Smith, Yanchuan Sim and Amber E. Boydston. 2020. "Etch-a-Sketching: Evaluating the Post-Primary Rhetorical Moderation Hypothesis." *American Politics Research* 48(1):99–131.
- Adams, James and Samuel Merrill. 2003. "Voter Turnout and Candidate Strategies in American Elections." *The Journal of Politics* 65(1):161–189.
- Adams, James and Samuel Merrill. 2014. "Candidates' Policy Strategies in Primary Elections: Does Strategic Voting by the Primary Electorate Matter?" *Public Choice* 160(1):7–24.
- Aldrich, John H. 1983. "A Downsian Spatial Model with Party Activism." *American Political Science Review* 77(4):974–990.
- Ansolabehere, Stephen, James M. Snyder and Charles Stewart. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45(1):136–159.
- Bailey, Michael. 2023. "Measuring Candidate Ideology from Congressional Tweets and Websites." *Political Analysis* .
- Bailey, Michael and Benjamin F. Reese. 2025. "Ideological Moderation and Success in U.S. Elections, 2020–2022." Working paper. SSRN. <https://doi.org/10.2139/ssrn.5169898>.
- Bawn, Kathleen, Martin Cohen, David Karol, Seth Masket, Hans Noel and John Zaller. 2012. "A Theory of Political Parties: Groups, Policy Demands and Nominations in American Politics." *Perspectives on Politics* 10(3):571–597.
- Bejarano, Christina E. 2013. *The Latina Advantage: Gender, Race, and Political Success*. Austin, TX: University of Texas Press.

- Brady, David W., Hahrie Han and Jeremy C. Pope. 2007. "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" *Legislative Studies Quarterly* 32(1):79–105.
- Brewer, Mark D. and Jeffrey M. Stonecash. 2001. "Class, Race Issues, and Declining White Support for the Democratic Party in the South." *Political Behavior* 23(2):131–155.
- Brown, Nadia E. 2014. *Sisters in the Statehouse: Black Women and Legislative Decision Making*. New York: Oxford University Press.
- Brown, Nadia E. and Danielle Casarez Lemi. 2021. *Sister Style: The Politics of Appearance for Black Women Political Elites*. New York: Oxford University Press.
- Burden, Barry C. 2001. "The Polarizing Effects of Congressional Primaries." *Representation and Electoral System* .
- Burnham, Colin and et al. 2025. "Zero-Shot Classification for Political Text Analysis." *Political Analysis* .
- Calonico, Sebastian, Matias D. Cattaneo and Rocío Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.
- Carew, Jessica D. H. 2016. "Stereotyping of Black Women and How It Affects Them in an Electoral Context." *National Political Science Review* .
- Center for American Women and Politics. 2025. "Women in State Legislatures 2025." **URL:** <https://cawp.rutgers.edu>
- Chen, Philip and Ashley Sorensen. 2025. "Raced/Gendered Electability: Support, Donations, and Democratic Double Standards for Black Women Candidates." *Political Research Quarterly* 78(3):942–956.

- Chirco, Patrizia and Tonya M. Buchanan. 2022. "Dark Faces in White Spaces: The Effects of Skin Tone, Race/Ethnicity, and Intergroup Preferences on Candidate Evaluations." *Analyses of Social Issues and Public Policy* 22(1):427–447.
- Cowburn, Mike and Marius Sältzer. 2025. "Partisan Communication in Two-Stage Elections: The Effect of Primaries on Intra-Campaign Positional Shifts in Congressional Elections." *Political Science Research and Methods* 13(2):392–411.
- Crenshaw, Kimberlé. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum* 1989:139–167.
- Dowe, Pearl K. Ford. 2020. "Resisting Marginalization: Black Women's Political Ambition and Agency." *PS: Political Science & Politics* 53(4):697–702.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Fenno, Richard F. 2003. *Going Home: Black Representatives and Their Constituents*. Chicago, IL: University of Chicago Press.
- Freyaldenhoven, Simon, Christian Hansen and Jesse M. Shapiro. 2019. "Pre-event Trends in the Panel Event-Study Design." *American Economic Review* 109(9):3307–3338.
- Fulton, Sarah A. and Sarah Allen Gershon. 2018. "Too Liberal to Win? Race and Voter Perceptions of Candidate Ideology." *American Politics Research* 46(5):909–939.
- Gay, Claudine and Katherine Tate. 1998. "Doubly Bound: The Impact of Gender and Race on the Politics of Black Women." *Political Psychology* 19(1):169–184.
- Gillespie, Andra. 2012. *The New Black Politician: Cory Booker, Newark, and Post-Racial America*. New York: New York University Press.
- Griswold, Alex. 2023. "Parker Suggests Support for Stop-and-Frisk Tactics." Philadelphia Inquirer.

- Hancock, Ange-Marie. 2004. *The Politics of Disgust: The Public Identity of the Welfare Queen*. New York: New York University Press.
- Hancock, Ange-Marie. 2007. "When Multiplication Doesn't Equal Quick Addition: Examining Intersectionality as a Research Paradigm." *Perspectives on Politics* 5(1):63–79.
- Harris-Perry, Melissa V. 2011. *Sister Citizen: Shame, Stereotypes, and Black Women in America*. New Haven, CT: Yale University Press.
- Heseltine, Michael and Bernhard Clemm von Hohenberg. 2024. "Zero-Shot Ideology Scaling of Text Using Natural Language Inference." *Political Analysis* .
- Hochschild, Jennifer and Vesla Weaver. 2007. "The Skin Color Paradox and the American Racial Order." *Social Forces* 86(2):643–670.
- Holman, Mirya R. and Monica C. Schneider. 2018. "Gender, Race, and Political Ambition: How Intersectionality and Frames Influence Interest in Political Office." *Politics, Groups, and Identities* 6(2):264–280.
- Huddy, Leonie and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37(1):119–147.
- Hummel, Patrick. 2010. "Flip-Flopping from Primaries to General Elections." *Journal of Public Economics* 94(11–12):1020–1027.
- King, David C. and Richard E. Matland. 2003. "Sex and the Grand Old Party: An Experimental Investigation of the Effect of Candidate Sex on Support for a Republican Candidate." *American Politics Research* 31(6):595–612.
- Koch, Jeffrey W. 1999. "Candidate Gender and Assessments of Senate Candidates." *Social Science Quarterly* 80(1):84–96.
- Landis, J. Richard and Gary G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33(1):159–174.

- Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2):281–355.
- Lemi, Danielle Casarez and Nadia E. Brown. 2019. "Melanin and Curls: Evaluation of Black Women Candidates." *Journal of Race, Ethnicity, and Politics* 4(2):259–296.
- Lemi, Danielle Casarez and Nadia E. Brown. 2021. "The Political Implications of Colorism Are Gendered." *American Politics Research* 49(5):514–524.
- Lerman, Amy E. and Meredith L. Sadin. 2016. "Stereotyping or Discrimination? The Role of Race in Shaping Voter Evaluations of State Legislatures." *Political Psychology* 37(5):621–635.
- Lewis, Mike, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov and Luke Zettlemoyer. 2020. "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension." *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* pp. 7871–7880.
- Mangum, Maruice. 2013. "Racial Heterogeneity and Ideological Representation in the U.S. House of Representatives." *Journal of Black Studies* 44(2):178–197.
- McCormick, Joseph P. and Charles E. Jones. 1993. The Conceptualization of Deracialization: Thinking Through the Dilemma. In *Dilemmas of Black Politics: Issues of Leadership and Strategy*, ed. Georgia Persons. New York: HarperCollins pp. 66–84.
- McDermott, Monika L. 1998. "Race and Gender Cues in Low-Information Elections." *Political Research Quarterly* 51(4):895–918.
- Miller, Gary and Norman Schofield. 2008. "The Transformation of the Republican and Democratic Party Coalitions in the United States." *Perspectives on Politics* 6(3):433–450.

- Nickels, Samantha. 2023. "Cherelle Parker's Record on Public Safety." Philadelphia Inquirer.
- Nokken, Timothy P. and Keith T. Poole. 2004. "Congressional Party Defection in American History." *Legislative Studies Quarterly* 29(4):545–568.
- Orey, Byron D' Andra and Jinhao Zhang. 2019. "Melanated Millennials and the Politics of Black Hair." *Social Science Quarterly* 100(6):2458–2476.
- Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357–384.
- Santia, Martina, Nichole M. Bauer and Sylvia Gonzalez. 2025. "Harnessing the Power of Emotions: How Women of Color Use Emotional Appeals in Their Campaign Messages." *Political Research Quarterly* 78(1):401–414.
- Schneider, Monica C. and Angela L. Bos. 2011. "An Exploration of the Content of Stereotypes of Black Politicians." *Political Psychology* 32(2):205–233.
- Schneider, Monica C. and Angela L. Bos. 2016. "The Interplay of Candidate Party and Gender in Evaluations of Political Candidates." *Journal of Women, Politics & Policy* 37(3):274–300.
- Shah, Paru, Jamil Scott and Eric Gonzalez Juenke. 2019. "Women of Color Candidates: Examining Emergence and Success in State Legislative Elections." *Politics, Groups, and Identities* 7(2):429–443.
- Sides, John, Michael Tesler and Lynn Vavreck. 2018. *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America*. Princeton, NJ: Princeton University Press.
- Sigelman, Carol K., Lee Sigelman, Barbara J. Walkosz and Michael Nitz. 1995. "Black

- Candidates, White Voters: Understanding Racial Bias in Political Perceptions.” *American Journal of Political Science* 39(1):243–265.
- Silva, Andrea and Carrie Skulley. 2019. “Always Running: Candidate Emergence among Women of Color.” *Political Research Quarterly* 72(2):342–354.
- Smith, Rogers M. 2004. “Identities, Interests, and the Future of Political Science.” *Perspectives on Politics* 2(2):301–312.
- Snipes, Alexandra. 2025. “Who’s Got the Best Shot? Race/Gender Stereotypes and Primary Elections.” *Political Research Quarterly* 78(4):1510–1529.
- Wasserman, David. 2025. “Cook Political Report: Partisan Voting Index.”
URL: <https://www.cookpolitical.com/pvi-map-and-district-list>
- WHYY. 2023. “Mayor Parker on Stop-and-Frisk: “There Is No Place for Unconstitutional” Tactics.”. WHYY News.
- Wittenbrink, Benjamin. 2022. “Analyzing Candidates’ Ideological Messaging Throughout the Electoral Cycle.” Honors thesis, Department of Economics, Stanford University. Advised by Matthew Gentzkow. <https://doi.org/10.25740/tg119qr1168>.
- Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz et al. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. pp. 38–45.