

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Zachary Baum

Department of Political Science

Rutgers, The State University of New Jersey

New Brunswick, NJ, USA

ORCID: 0009-0003-7374-684X

Email: *zb207@scarletmail.rutgers.edu*

Corresponding Author

Zachary Baum

Department of Political Science

Rutgers, The State University of New Jersey

89 George Street

New Brunswick, NJ 08901

United States

Author Note

The author declares no conflicts of interest.

Funding Statement

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

This research received support from the Institute for Humane Studies through the Humane Studies Fellowship. The funder had no role in the design of the study; the collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to submit the article for publication.

Data Availability Statement

The data, replication materials, experimental prompts, candidate profiles, and analysis code supporting the findings of this study are available from the corresponding author upon reasonable request.

Ethics Statement

This study did not involve human participants. The research analyzed outputs generated by publicly available large language models using standardized prompts. Institutional Review Board approval was therefore not required.

Acknowledgments

The author thanks colleagues in the Department of Political Science at Rutgers University for helpful discussions during the development of this project.

Author Contributions

Conceptualization: Zachary Baum

Methodology: Zachary Baum

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Investigation: Zachary Baum

Data Curation: Zachary Baum

Formal Analysis: Zachary Baum

Visualization: Zachary Baum

Writing – Original Draft: Zachary Baum

Writing – Review & Editing: Zachary Baum

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Abstract

Political leadership has long been evaluated through social expectations surrounding authority, professionalism, electability, and legitimacy. As large language models (LLMs) become increasingly integrated into political communication, campaigns, journalism, and information search, understanding how these systems evaluate political candidates has become an important scholarly question. Despite growing research on algorithmic bias, comparatively little attention has been devoted to AI-generated assessments of political candidates or whether those assessments vary systematically across candidate identities.

This study examines whether commercially available generative AI systems produce systematically different evaluations of otherwise comparable political candidates when sexual orientation, gender identity, and leadership presentation vary while substantive qualifications remain constant. Using a standardized vignette experiment, eighteen candidate profiles were evaluated by six widely used large language models across repeated independent sessions, yielding 540 AI-generated candidate evaluations. The analysis compares assessments of electability, leadership, competence, political strength, and overall candidate quality across candidate conditions.

The findings indicate that AI-generated differences were concentrated primarily in assessments of electability and political strength rather than competence. Across models, LGBTQ candidates generally received modestly lower evaluations of political viability than otherwise comparable non-LGBTQ candidates, while candidates described using conventionally masculine leadership characteristics received substantially higher evaluations than candidates

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

emphasizing collaborative or nontraditional leadership styles. The magnitude of these differences varied across AI systems.

This article introduces the concept of algorithmic electability to describe systematic variation in AI-generated assessments of candidate viability under otherwise equivalent informational conditions. Rather than drawing conclusions about the origins of these differences or their downstream effects on political behavior, the study documents AI-generated candidate evaluations as an emerging object of inquiry and extends research on candidate evaluation and political communication into the domain of generative artificial intelligence.

Keywords: generative artificial intelligence; LGBTQ candidates; algorithmic bias; electability; descriptive representation; candidate evaluation; gender nonconformity; political behavior; AI and democracy; large language models; political communication

Introduction

Generative artificial intelligence has rapidly become embedded within contemporary political communication. Large language models (LLMs) are increasingly used to summarize political information, evaluate candidates, draft campaign materials, synthesize news coverage, and answer questions regarding elections and public affairs. Rather than functioning solely as passive repositories of information, these systems generate evaluative judgments that increasingly shape how users encounter and interpret political actors.

Political communication scholars have long argued that information intermediaries—including journalists, search engines, and digital platforms—play an important role in structuring political knowledge and influencing perceptions of candidates (Kreiss &

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

McGregor, 2019). As generative AI systems become integrated into search engines, productivity software, campaign workflows, and public-facing information platforms, understanding the content of AI-generated political evaluations has become an increasingly important scholarly question.

Existing research on algorithmic bias has primarily focused on domains such as employment, criminal justice, facial recognition, lending, and content moderation (Benjamin, 2019; Noble, 2018; Crawford, 2021). Comparatively little attention has been devoted to how generative AI systems evaluate political candidates or whether those evaluations vary systematically across candidate identities. This omission is increasingly consequential given the expanding role of AI-generated information within democratic communication environments.

The present study therefore examines AI-generated candidate evaluations as an object of political inquiry. Rather than asking whether artificial intelligence influences elections directly, it investigates whether contemporary large language models generate systematically different evaluations of otherwise comparable political candidates under standardized informational conditions.

Candidate Evaluation and Electability Heuristics

Political scientists have long distinguished between judgments of candidate competence and judgments of candidate electability. Research documents that voters, donors, journalists, campaign professionals, and party elites frequently evaluate candidates not only according to

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

their qualifications but also according to beliefs about whether those candidates are capable of winning elections (Lawless & Fox, 2010; Sanbonmatsu, 2002; Brooks, 2013).

These electability heuristics often operate independently of evaluations of competence. Individuals may personally regard a candidate as highly qualified while simultaneously expressing doubts about the candidate's electoral prospects because they believe other voters may respond negatively to particular social identities or demographic characteristics. Such strategic evaluations influence candidate recruitment, fundraising, endorsements, media coverage, and political ambition.

Research on women candidates illustrates this distinction particularly well. Although explicit evaluations of competence have become increasingly equal over time, women frequently continue to encounter skepticism regarding their electoral viability (Brooks, 2013). Similar patterns have been documented for racial minority candidates and LGBTQ candidates, where strategic concerns regarding voter acceptance frequently persist despite favorable assessments of professional qualifications.

Although these electability heuristics have been studied extensively among human decision makers, little research has examined whether comparable patterns emerge in AI-generated political evaluations. If generative AI systems produce similar distinctions between competence and electability, they may reproduce longstanding political heuristics within digital information environments.

Gendered Leadership Prototypes

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Research on political leadership further suggests that evaluations of electoral viability are shaped by culturally embedded expectations regarding legitimate leadership. Political leadership has historically been associated with characteristics such as decisiveness, authority, confidence, rationality, toughness, and institutional professionalism, traits that have frequently been socially coded as masculine (Eagly & Karau, 2002).

Role congruity theory argues that individuals whose identities or leadership styles depart from these expectations may encounter evaluative penalties because they violate dominant leadership prototypes rather than because they lack competence (Eagly & Karau, 2002). Political behavior research similarly shows that voters often distinguish between candidate qualifications and candidate "fit" with conventional expectations regarding political leadership (Huddy & Terkildsen, 1993).

These leadership schemas are particularly relevant for understanding evaluations of transgender and gender-nonconforming candidates. Departures from conventional gender presentation may challenge implicit assumptions regarding political legitimacy even when objective qualifications remain equivalent. Consequently, AI-generated evaluations may reproduce differences associated not only with candidate identity but also with the leadership characteristics emphasized within candidate biographies.

LGBTQ Candidates and Political Representation

Research examining LGBTQ candidates clarifies that sexual orientation and gender identity can shape political evaluations in complex ways. LGBTQ candidates have historically

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

confronted barriers related to visibility, perceived acceptability, fundraising, media coverage, and assumptions regarding electoral competitiveness (Haider-Markel, 2010; Reynolds, 2013).

Importantly, these barriers frequently reflect strategic judgments rather than direct assessments of competence. Voters and political elites may view LGBTQ candidates favorably while simultaneously questioning whether other voters will support them electorally. Such distinctions mirror broader research demonstrating that perceptions of electability often differ from perceptions of candidate quality.

These findings suggest that candidate identity may influence evaluations through assessments of political viability rather than professional capability. Whether contemporary AI systems generate comparable distinctions remains an open empirical question.

Algorithmic Bias and Political Evaluation

Scholarship on algorithmic bias has demonstrated that computational systems can reproduce patterns of inequality present within the data, institutions, and social environments from which they are developed (Benjamin, 2019; Noble, 2018; Crawford, 2021). Existing studies have documented unequal outcomes across domains including employment, lending, criminal justice, facial recognition, and online search.

Generative AI presents a distinct challenge because large language models do not merely retrieve information; they generate evaluative narratives and recommendations. Consequently, these systems may produce different assessments of political candidates even when presented with otherwise equivalent information.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Importantly, the present study does not assume that observed differences arise from any particular feature of model development. Contemporary AI systems differ in architecture, alignment procedures, safety mechanisms, and training approaches, but the proprietary nature of these systems makes causal attribution impossible. Accordingly, the analysis focuses on documenting variation in AI-generated evaluations rather than explaining the mechanisms responsible for those differences.

Theoretical Contribution

This study brings together two largely separate literatures. Political science has developed extensive scholarship on candidate evaluation, electability heuristics, and descriptive representation, while computer science and information studies have documented algorithmic bias across numerous policy domains. Comparatively little research has examined whether generative AI systems themselves generate systematically different evaluations of political candidates.

The concept of algorithmic electability introduced here refers to systematic variation in AI-generated assessments of candidate viability under otherwise equivalent informational conditions. Rather than treating AI systems as neutral technologies or assuming that they inevitably reproduce political inequality, the concept directs attention to the content of AI-generated evaluations themselves and to whether those evaluations differ across candidate identities.

By examining standardized candidate profiles across multiple commercially available AI systems, the present study evaluates whether AI-generated assessments distinguish between

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

competence and electability, whether such differences vary across candidate identities, and whether those patterns differ across models. In doing so, it extends scholarship on candidate evaluation into the rapidly evolving domain of generative artificial intelligence while identifying AI-generated political evaluation as a new object of inquiry within political communication and representation research.

This article makes three principal contributions to scholarship on political communication, candidate evaluation, and artificial intelligence.

First, it introduces the concept of algorithmic electability to describe systematic variation in AI-generated assessments of candidate viability under otherwise equivalent informational conditions. Existing research has extensively examined how voters, party elites, journalists, and donors evaluate political candidates, while scholarship on algorithmic bias has focused primarily on domains such as employment, criminal justice, lending, and content moderation. By conceptualizing AI-generated candidate evaluations as a distinct form of political judgment, this study extends theories of electability into the rapidly expanding domain of generative artificial intelligence.

Second, the study contributes to research on candidate evaluation by demonstrating that AI-generated differences are concentrated primarily in assessments of electability and political strength rather than competence. This distinction mirrors longstanding findings in political behavior research showing that candidates may be viewed as professionally qualified while simultaneously being perceived as less electorally viable because of social identities or departures from conventional leadership expectations. Examining whether similar distinctions

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

emerge in AI-generated evaluations broadens existing theories of electability beyond human decision-makers to include increasingly influential computational systems.

Third, by comparing six commercially available large language models using standardized candidate vignettes, this study documents meaningful variation across contemporary AI systems. Rather than treating generative AI as a single, homogeneous technology, the analysis demonstrates that AI-generated political evaluations differ across models under identical informational conditions. Although the present study does not identify the mechanisms responsible for those differences, it establishes AI-generated candidate evaluation as an important new object of inquiry and provides a foundation for future research examining how generative AI may shape political communication, candidate assessment, and democratic representation.

Taken together, these contributions position AI-generated candidate evaluation as a theoretically meaningful extension of the candidate evaluation literature while introducing algorithmic electability as a framework for studying how large language models assess political viability in contemporary democratic politics.

Expectations

H1: Candidates whose leadership presentation departs from conventional masculine political norms will receive lower evaluations of electability and leadership than otherwise comparable candidates.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

H2: LGBTQ candidates will receive lower electability evaluations than otherwise comparable non-LGBTQ candidates.

H3: Differences associated with candidate identity will be larger for electability than for competence evaluations.

H4: The magnitude of identity-related evaluation differences will vary across AI systems.

Methods

This study employed a standardized vignette-based evaluation design to examine whether contemporary large language models (LLMs) generate systematically different evaluations of political candidates when candidate identity cues vary while substantive qualifications remain constant. Rather than measuring the attitudes or behaviors of human respondents, the unit of analysis was the evaluative output produced by commercially available generative AI systems in response to identical candidate profiles.

The research design isolates the effect of identity-related cues by holding candidates' qualifications, professional experience, policy positions, educational background, endorsements, and community involvement constant while varying selected characteristics related to sexual orientation, transgender identity, gender presentation, and leadership style. This design permits comparison of AI-generated evaluations under standardized informational conditions.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Because the objective is to characterize AI-generated evaluations rather than estimate public opinion, the study should be understood as an evaluation of model behavior rather than a behavioral experiment involving human participants.

Candidate Vignette Construction

Candidate profiles were designed to resemble realistic candidates for state legislative office in the United States. Each vignette included information regarding:

- educational attainment;
- occupational experience;
- community involvement;
- prior public service;
- policy priorities;
- endorsements;
- family background; and
- campaign messaging.

Across all conditions, these substantive qualifications remained constant. Experimental manipulations altered only identity-related characteristics or leadership presentation while preserving the underlying candidate profile.

Identity manipulations included references to:

- same-sex spouses,
- LGBTQ community involvement,

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

- transgender identity,
- gender ambiguity,
- and nontraditional gender presentation.

Leadership manipulations altered descriptive language associated with political leadership styles. Traditional leadership conditions emphasized characteristics such as decisiveness, pragmatism, confidence, discipline, and institutional professionalism. Nontraditional leadership conditions emphasized collaboration, empathy, emotional openness, expressiveness, and participatory leadership.

A total of eighteen unique candidate vignettes were developed to capture the experimental conditions examined in the study.

AI Systems Evaluated

The study evaluated responses generated by six commercially available large language models:

- ChatGPT
- Claude
- Gemini
- Grok
- Perplexity
- DeepSeek

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

These systems were selected because they represent widely used publicly accessible generative AI platforms that differ in training procedures, alignment strategies, and model architecture.

The purpose of including multiple models was not to identify the causes of model-specific differences but rather to evaluate whether AI-generated political assessments vary across contemporary commercially available systems.

Prompt Standardization

Each AI platform received identical prompts and identical candidate vignettes.

Prompts instructed the model to evaluate each candidate across multiple dimensions including:

- electability;
- leadership ability;
- competence;
- political strength; and
- overall candidate quality.

To minimize prompt-induced variation, wording remained constant across all AI platforms and all candidate conditions. Candidate identity cues represented the only intentional source of variation across treatment conditions.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

The complete prompt text is reproduced in Appendix A to facilitate replication and transparency.

Evaluation Procedure

Each candidate vignette was evaluated five independent times within each AI platform.

Independent evaluations were conducted using separate chat sessions to reduce the influence of conversational memory and stochastic response variation. Repeated evaluations were collected because contemporary LLMs generate probabilistic outputs that may vary across identical prompts.

Across eighteen candidate profiles, six AI systems, and five independent evaluations per condition, the final dataset consisted of 540 AI-generated evaluations.

Repeated evaluations were retained rather than averaged to characterize within-model variability and to permit estimation of uncertainty across repeated model outputs.

Outcome Coding

AI-generated responses were converted into quantitative outcome measures using a standardized coding protocol.

Most AI systems generated explicit numerical ratings. When responses included accompanying narrative explanations, numerical scores were extracted directly while qualitative text was retained for verification and interpretation.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Because the principal dependent variables consisted of explicit numerical evaluations generated directly by the models, intercoder reliability was not applicable for these quantitative measures. Examples of representative AI outputs are included in Appendix B.

Statistical Analysis

The analysis proceeded in several stages.

First, descriptive statistics summarized average evaluations across experimental conditions.

Second, Welch two-sample t-tests compared mean evaluations between LGBTQ and non-LGBTQ candidate profiles.

Third, ordinary least squares regression models with heteroskedasticity-robust standard errors estimated the association between candidate identity cues and AI-generated evaluations. Interaction terms between candidate identity and AI platform were subsequently included to assess whether observed evaluation differences varied across systems. Because repeated evaluations were generated within the same AI platforms and candidate vignettes, observations may not be fully independent. The regression estimates should therefore be interpreted as descriptive associations rather than definitive causal or population-level effects.

Interaction models subsequently examined whether estimated identity-related evaluation differences varied across AI platforms.

Robustness Checks

Several robustness analyses were conducted to evaluate the stability of the findings.

First, analyses were replicated separately for each AI platform.

Second, interaction models assessed whether estimated identity effects differed significantly across models.

Third, results were compared across electability, competence, leadership, and political strength outcomes to determine whether observed differences reflected generalized candidate evaluations or were concentrated in specific dimensions of political assessment.

Across specifications, the central pattern remained consistent: identity-related differences emerged primarily in electability and political strength rather than competence.

Reproducibility and Limitations

Generative AI systems are proprietary technologies that evolve over time through updates to model architecture, alignment procedures, safety systems, and training data. Consequently, identical prompts administered at different points in time may produce different outputs.

The objective of this study is therefore not to establish immutable properties of any particular model but rather to characterize the behavior of contemporary commercially available LLMs during the study period.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Interaction analyses indicated meaningful variation across AI systems. ChatGPT generated the largest average electability differences between LGBTQ and non-LGBTQ candidates, whereas Gemini, DeepSeek, Claude, and Perplexity exhibited substantially smaller differences.

These findings demonstrate that AI-generated candidate evaluations were not identical across contemporary large language models. Because the present study does not evaluate model architecture or training procedures directly, however, the analysis cannot determine why these differences emerged.

Accordingly, the observed model-level variation should be interpreted descriptively rather than as evidence regarding the causes of algorithmic differences.

Table 1
Experimental Candidate Conditions

Condition	Experimental Manipulation
Control	Opposite-sex spouse cue
LGBTQ	Same-sex spouse cue
Transgender	Explicit transgender identity
Gender Ambiguous	Neutralized gender markers
Nontraditional Leadership	Expressive, collaborative, or unconventional leadership cues
Traditional Leadership	Disciplined, pragmatic, conventionally authoritative leadership cues

Several limitations should be considered when interpreting these findings.

First, this study evaluates AI-generated candidate evaluations rather than human political behavior. Although large language models are increasingly used to summarize political

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

information, generate strategic recommendations, and evaluate candidates, the present design does not examine whether AI-generated evaluations influence voters, donors, journalists, campaign professionals, political consultants, or party elites. Consequently, the findings should not be interpreted as evidence that observed differences translate directly into electoral outcomes or political decision making. Rather, the study documents systematic variation in AI-generated evaluations under standardized conditions.

Second, the study evaluates model outputs rather than model architecture. Contemporary large language models are proprietary systems whose training corpora, alignment procedures, reinforcement learning methods, moderation systems, and parameter configurations are not publicly available in sufficient detail to permit causal attribution. Although meaningful differences were observed across AI platforms, the present analysis cannot determine whether those differences arise from training data, alignment strategies, reinforcement learning, safety mechanisms, model architecture, or other aspects of system development. These mechanisms remain important questions for future research.

Third, large language models are dynamic systems. AI developers regularly update model weights, alignment procedures, safety policies, and deployment settings. As a result, identical prompts administered at different points in time may produce different outputs. The findings reported here therefore characterize model behavior during the period in which the data were collected rather than establishing permanent properties of any individual AI system.

Fourth, the experimental design necessarily simplifies real-world candidate evaluation. Candidate profiles were intentionally standardized to isolate the effects of identity-related cues

while holding qualifications constant. Actual political campaigns involve substantially richer information environments, including partisan affiliation, incumbency, campaign resources, media coverage, issue positions, demographic context, and strategic interactions that are not represented in vignette-based evaluations. Accordingly, the observed differences should be interpreted within the controlled conditions of the experimental design rather than as complete representations of electoral politics.

Fifth, the study focuses on a limited set of candidate identities and leadership presentations. Although the analyses examine sexual orientation, transgender identity, gender presentation, and leadership style, they do not evaluate many other politically relevant identities or their intersections, including race, ethnicity, disability, religion, age, or socioeconomic background. Future research should examine how multiple marginalized identities interact within AI-generated political evaluations.

Finally, the present study documents systematic differences in AI-generated evaluations but cannot determine whether those differences reflect broader societal stereotypes, patterns present in publicly available political discourse, design choices made by AI developers, or other mechanisms. Identifying the origins of these observed differences will require complementary research combining computational analyses, model auditing, transparency initiatives, and behavioral experiments involving human participants.

Taken together, these limitations do not diminish the contribution of the present study. Rather, they clarify the scope of the inferences that can reasonably be drawn. The findings provide evidence that contemporary large language models can generate systematically different

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

evaluations of otherwise comparable political candidates under standardized conditions, while leaving important questions regarding the mechanisms and downstream consequences of those differences for future research.

Results

Across all AI systems, evaluations consistently distinguished between candidate competence and candidate viability. While competence ratings differed little across identity conditions, electability and political strength exhibited more systematic variation. These patterns suggest that any observed differences in AI-generated evaluations were concentrated primarily in assessments related to electoral competitiveness rather than perceptions of professional qualifications.

Average electability scores were modestly higher for non-LGBTQ candidates ($M = 7.26$) than for otherwise comparable LGBTQ candidates ($M = 7.01$), representing an average difference of approximately 0.25 points on the ten-point evaluation scale. Political strength exhibited a similar pattern, with non-LGBTQ candidates receiving average scores of 7.43 compared to 7.25 for LGBTQ candidates. By contrast, competence evaluations differed negligibly across conditions, with both groups receiving average ratings of approximately 7.5.

Although the observed differences in electability and political strength were relatively modest in absolute magnitude, they appeared consistently across multiple AI systems and repeated candidate evaluations.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Welch two-sample t-tests indicated that political strength evaluations differed significantly across identity conditions. Non-LGBTQ candidates received higher average political strength evaluations than LGBTQ candidates ($t = 2.73$, $p = .0066$). The estimated difference of approximately 0.19 points corresponds to a modest but statistically reliable reduction in AI-generated assessments of political strength.

By contrast, differences in competence evaluations were substantively negligible and statistically indistinguishable from zero. This distinction between competence and viability represents one of the central empirical findings of the study, suggesting that AI systems did not characterize LGBTQ candidates as less qualified but instead evaluated them somewhat less favorably on dimensions associated with electoral success.

Gendered and Identity-Based Candidate Penalties

The overall pattern across analyses was remarkably consistent. Differences in AI-generated evaluations appeared primarily in assessments related to electability, leadership, and political strength, whereas competence evaluations remained comparatively stable across candidate identities.

Rather than demonstrating generalized differences across all dimensions of candidate evaluation, the findings suggest that AI-generated variation was concentrated in judgments related to perceived political viability. This distinction provides the central empirical contribution of the study.

Transgender Candidate Evaluations

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Candidate profiles explicitly identifying transgender identity exhibited lower average electability evaluations than otherwise equivalent comparison profiles. Across models, the average electability score declined from 7.30 to 6.85, representing an estimated reduction of approximately 0.45 points on the ten-point evaluation scale.

Importantly, this pattern was not mirrored in competence evaluations. Average competence ratings remained nearly identical across conditions, differing by only 0.06 points. Political strength also declined modestly for transgender candidates relative to comparison profiles.

These findings suggest that AI-generated differences associated with transgender identity were concentrated primarily in assessments of political viability rather than professional capability. While the present study cannot determine the mechanisms responsible for these differences, the observed pattern is consistent with longstanding distinctions in political behavior research between judgments of competence and judgments of electability.

Gendered Leadership Norms and Political Legitimacy

The analysis also examined whether departures from conventional gender presentation influenced AI-generated political evaluations. Candidate profiles employing gender-neutral or ambiguous identity cues produced comparatively modest differences in evaluation, suggesting that ambiguity alone did not consistently trigger substantial penalties across models.

The clearest evidence of gendered algorithmic bias emerged in the leadership presentation experiment. Candidates framed through traditionally coded leadership

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

descriptors, including discipline, pragmatism, confidence, and conventional leadership suitability, received substantially more favorable evaluations than candidates framed through expressive, emotionally open, collaborative, and unconventional leadership characteristics.

The traditional leadership prototype condition received an average electability score of 7.68, compared to just 6.74 for the nontraditional leadership condition, representing a difference of nearly one full point on the ten-point evaluation scale. Similar disparities emerged in other outcome measures. Traditional leadership candidates received an average competence score of 7.71, compared to 7.35 in the nontraditional condition, while political strength evaluations declined from 7.70 to 7.06.

The leadership presentation manipulation produced the largest observed differences in the study. Candidate profiles emphasizing conventionally masculine leadership characteristics—including decisiveness, discipline, pragmatism, and confidence—received consistently higher evaluations than otherwise comparable candidates emphasizing collaborative, emotionally expressive, or nontraditional leadership styles.

Average electability differed by approximately 0.94 points across these conditions, substantially larger than the differences associated with LGBTQ identity alone. Similar patterns emerged for political strength and leadership evaluations, while competence differences remained comparatively modest.

These findings are consistent with the possibility that AI-generated evaluations privilege leadership characteristics historically associated with conventional political leadership. Because

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

the present study evaluates model outputs rather than underlying model architecture, however, it cannot determine why these differences emerged.

Regression Analysis

Regression analyses produced results consistent with the descriptive comparisons. Across specifications, LGBTQ identity was associated with lower average electability evaluations after accounting for other variables included in the model.

The estimated coefficient indicated an average reduction of approximately one-quarter of a point on the ten-point electability scale ($b = -0.248$, $p = .0015$). Although statistically reliable, the substantive magnitude of this effect was modest.

No statistically significant relationship emerged between LGBTQ identity and competence evaluations ($b = 0.003$, $p = .968$), reinforcing the broader descriptive finding that AI systems differentiated more strongly on assessments of political viability than professional capability.

Taken together, these findings indicate that observed identity-related differences were concentrated in electability-related outcomes rather than generalized across all dimensions of candidate evaluation.

Interaction Effects Across AI Systems

Interaction models further demonstrated substantial variation across AI systems. The magnitude of the LGBTQ electability penalty differed significantly depending on the model

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

generating the evaluation. ChatGPT exhibited the largest apparent electability penalty for LGBTQ candidates, while other systems, including Claude, DeepSeek, Gemini, and Perplexity, showed substantially smaller penalties or near-null effects after accounting for interaction terms.

These results indicate that AI-generated evaluations are not uniform across contemporary large language models. Under identical candidate profiles and standardized prompts, some models generated larger differences in electability evaluations than others. The present study, however, does not identify the mechanisms responsible for these differences. Because the analysis does not examine model architecture, training corpora, alignment procedures, or reinforcement learning directly, it cannot determine why these models produced different outputs. Accordingly, the observed variation should be interpreted descriptively rather than as evidence regarding the specific causes of algorithmic differences.

Figure 1. Electability Ratings by AI Model and LGBTQ Candidate Condition

Mean electability evaluations across six generative AI systems for LGBTQ and non-LGBTQ political candidates. Across most models, LGBTQ candidates received lower electability scores than otherwise comparable non-LGBTQ candidates, though the magnitude of the penalty varied substantially by AI system. ChatGPT and Grok exhibited the largest negative gaps in evaluations, while Gemini showed little to no apparent electability penalty. These findings suggest that algorithmic evaluations of political viability are not uniform across generative AI platforms and may reflect model-specific differences in training data, alignment procedures, or embedded political heuristics.

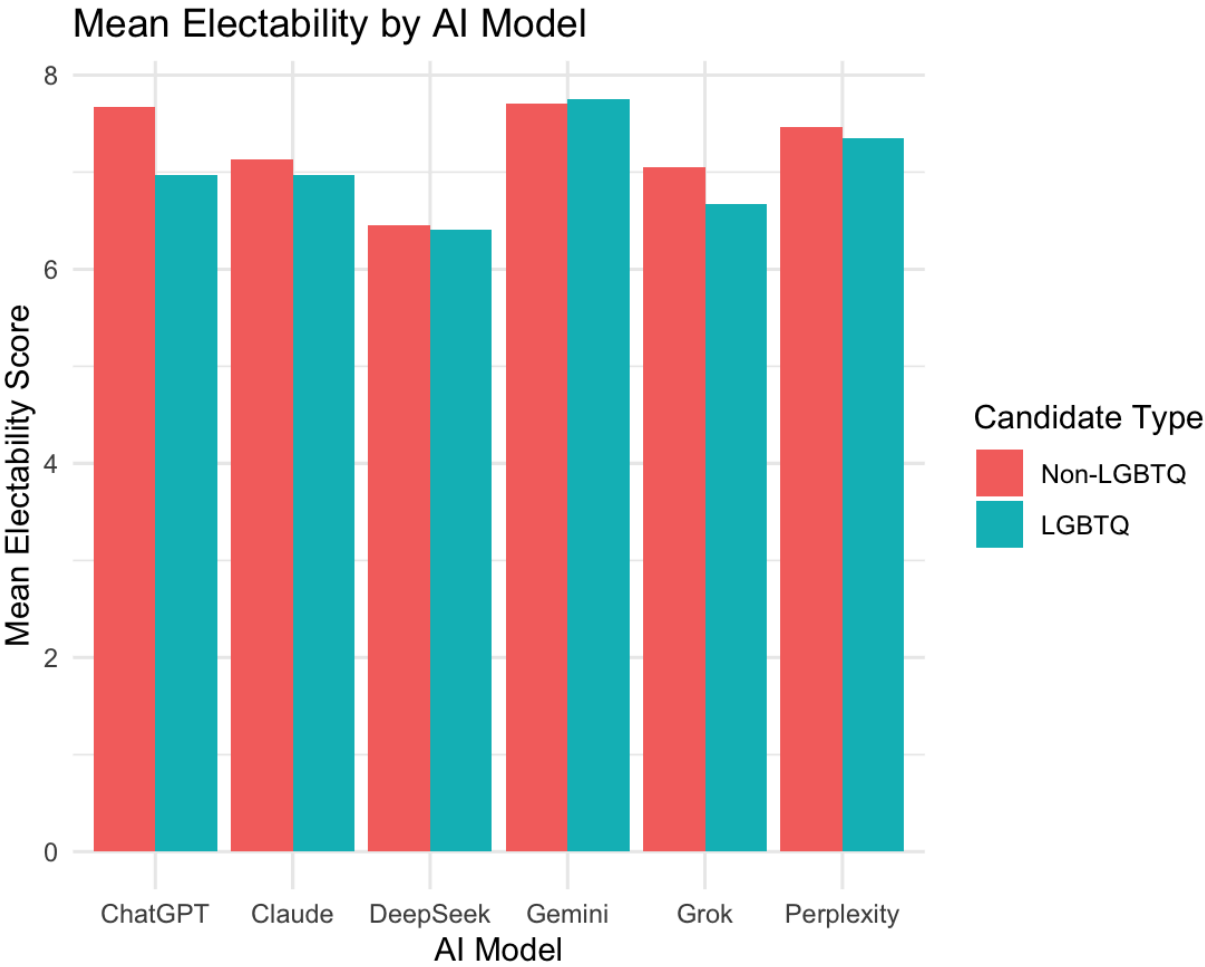


Figure 2. Strength Evaluations by LGBTQ Candidate Condition

Boxplots displaying the distribution of political strength evaluations for LGBTQ and non-LGBTQ candidates across all AI systems included in the study. Non-LGBTQ candidates generally received higher strength evaluations and exhibited a slightly higher median score compared to LGBTQ candidates. Although the distributions overlap substantially, LGBTQ candidates received lower average evaluations overall, consistent with the statistically significant differences identified in the regression and t-test analyses. The results suggest that generative AI systems may associate LGBTQ identity cues with reduced perceptions of political strength or viability even when candidate qualifications remain otherwise constant.

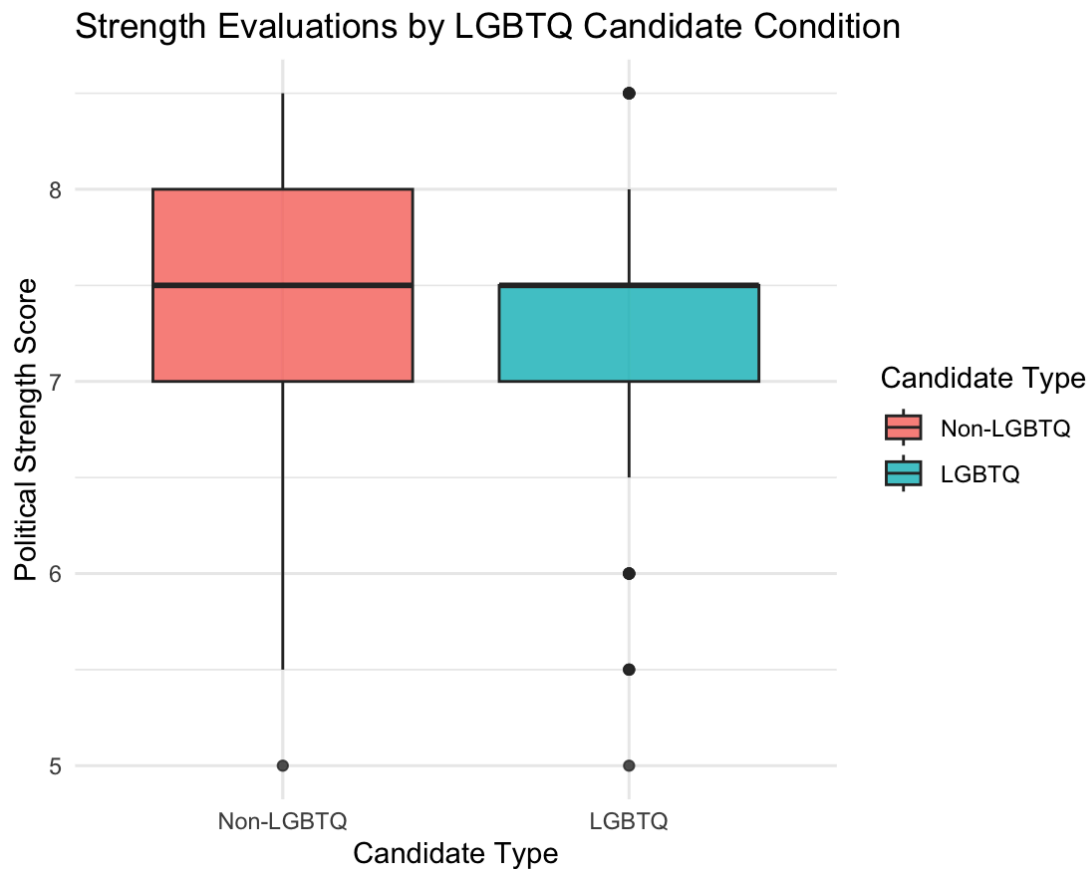


Figure 3. Electability Evaluations by LGBTQ Candidate Condition

Boxplots displaying the distribution of electability evaluations for LGBTQ and non-LGBTQ political candidates across all AI systems included in the study. Non-LGBTQ candidates generally received higher electability evaluations and exhibited a slightly higher central tendency than LGBTQ candidates, despite otherwise equivalent candidate qualifications across experimental conditions. While the distributions overlap considerably, LGBTQ candidates received consistently lower average electability scores overall, supporting the regression findings indicating a statistically significant electability penalty associated with LGBTQ identity cues. These findings suggest that generative AI systems may reproduce broader assumptions regarding political viability and perceived electoral competitiveness.

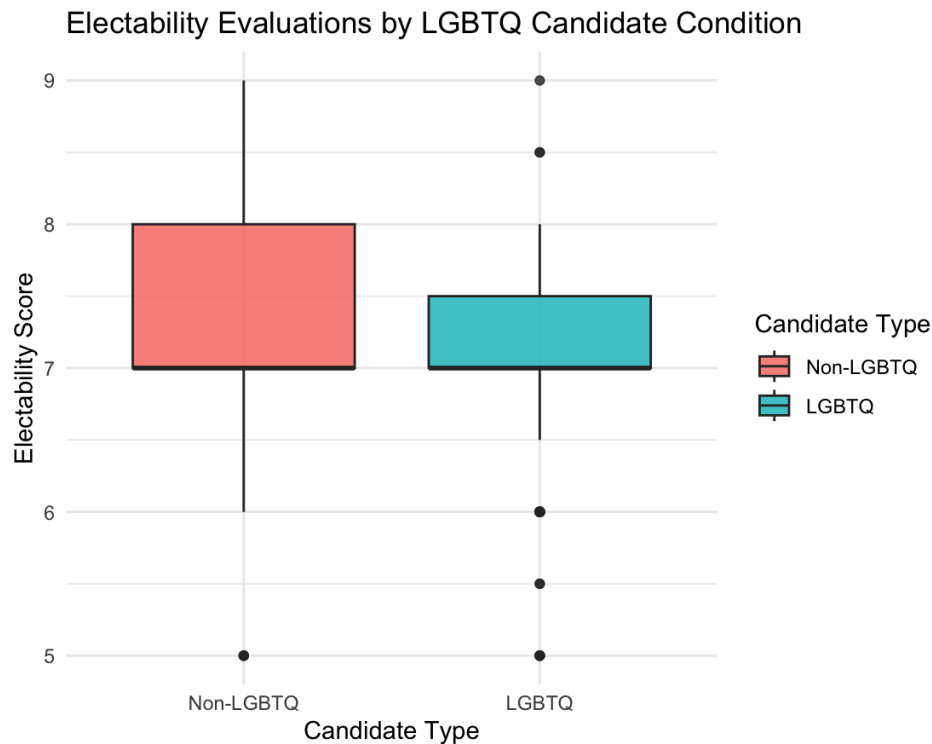
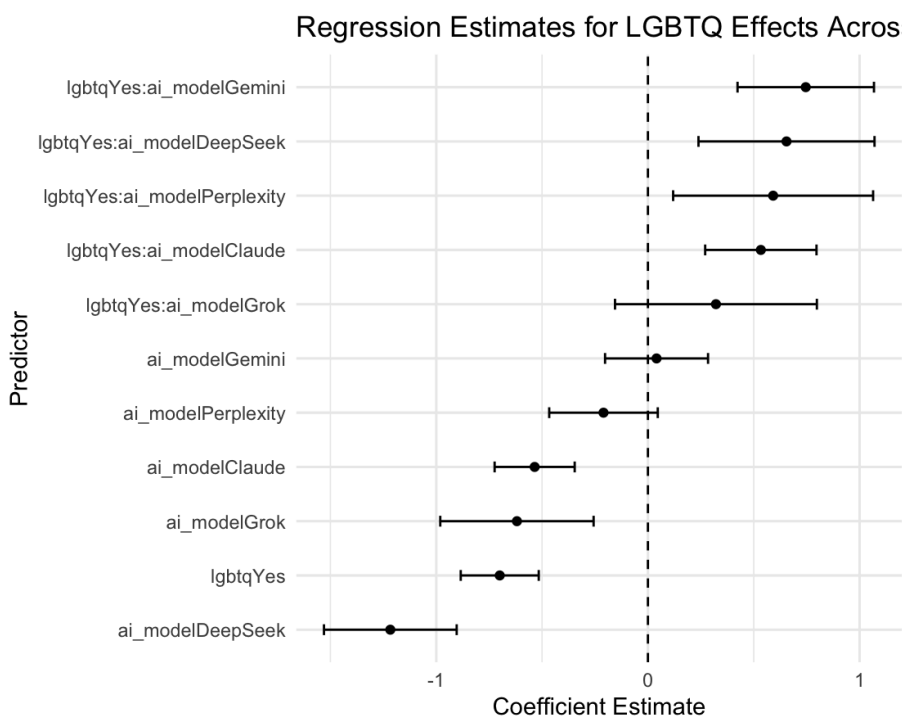


Figure 4. Regression Estimates for LGBTQ Effects Across AI Models

Coefficient plot displaying regression estimates and 95% confidence intervals from interaction models examining the relationship between LGBTQ candidate identity and electability evaluations across generative AI systems. The dashed vertical line at zero represents no estimated effect. Negative coefficients indicate lower electability evaluations relative to the reference category, while positive interaction terms indicate reductions in the magnitude of the LGBTQ electability penalty within specific AI models relative to the baseline model. The results demonstrate substantial variation across AI systems in how LGBTQ candidate cues shape perceptions of political viability, suggesting that algorithmic evaluations of electability are model-dependent rather than uniform across generative AI platforms.



Conclusion

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

This study examined whether contemporary large language models generate systematically different evaluations of political candidates when identity cues and leadership presentation vary while substantive qualifications remain constant. Across six commercially available AI systems, the analyses identified a consistent distinction between evaluations of competence and evaluations of political viability. Although competence ratings differed little across candidate conditions, differences emerged more consistently in assessments of electability, leadership, and political strength.

The findings suggest that AI-generated candidate evaluations may distinguish between professional qualifications and perceived electoral viability in ways that are consistent with longstanding theories of electability heuristics and gendered leadership expectations. In particular, candidates emphasizing conventionally masculine leadership characteristics generally received more favorable evaluations than otherwise comparable candidates emphasizing collaborative or nontraditional leadership styles. Similarly, LGBTQ identity cues were associated with modest reductions in assessments of political viability across several outcomes, although the magnitude of these differences varied considerably across AI systems.

Equally important, the results demonstrate that AI-generated political evaluations are not uniform across contemporary large language models. Identical candidate profiles produced different patterns of evaluation across platforms, underscoring the importance of treating AI systems as distinct political information intermediaries rather than assuming that generative AI produces homogeneous outputs. Because the present study evaluates model behavior rather than model architecture, however, it cannot determine why these differences emerged or identify the specific mechanisms responsible for variation across systems.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

The primary contribution of this article is conceptual. By introducing the concept of algorithmic electability, the study extends scholarship on candidate evaluation into the rapidly evolving domain of generative artificial intelligence. Rather than assuming that AI systems are either neutral technologies or direct reproductions of existing political inequalities, the concept directs attention to AI-generated candidate evaluations themselves as a distinct phenomenon worthy of empirical investigation.

Several limitations remain. The analyses characterize AI-generated outputs under standardized experimental conditions rather than human political behavior, and they do not establish whether these evaluations influence voters, journalists, campaign professionals, donors, or party elites. Nor do they identify the mechanisms through which different AI systems generate varying evaluations. Future research should examine whether AI-generated candidate assessments shape political decision-making in real-world settings, explore how these patterns evolve as models are updated, and investigate whether similar dynamics emerge across additional political identities, institutional contexts, and electoral systems.

As generative AI becomes increasingly embedded within political communication, campaign strategy, and public information environments, understanding how these systems evaluate political candidates represents an important new area of inquiry. By documenting systematic differences in AI-generated candidate evaluations across standardized profiles, this study provides an empirical foundation for future research on the relationship between artificial intelligence, candidate evaluation, and democratic representation.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

As generative AI becomes increasingly embedded within search engines, campaign consulting, journalism, candidate recruitment, and voter information environments, AI-generated evaluations may become an increasingly important component of political information ecosystems. Understanding whether these systems systematically evaluate candidates differently therefore represents not only a question of algorithmic fairness but also an emerging question for democratic representation and political communication.

References

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989.

<https://doi.org/10.1177/1461444816676645>

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity Press.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.

<https://doi.org/10.1145/3442188.3445922>

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of Machine Learning Research*, 81, 149–159.

Brooks, D. J. (2013). *He runs, she runs: Why gender stereotypes do not harm women candidates*. Princeton University Press.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.

Carroll, S. J., & Fox, R. L. (2018). *Gender and elections: Shaping the future of American politics* (4th ed.). Cambridge University Press.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Crawford, K., & Paglen, T. (2021). Excavating AI: The politics of training sets for machine learning. *AI & Society*, 36(4), 1105–1116. <https://doi.org/10.1007/s00146-021-01162-8>

Doan, A. E., Loehr, A., & Miller, L. R. (2014). Formal rights and informal privileges for same-sex couples: Evidence from a national survey experiment. *American Politics Research*, 42(5), 815–838. <https://doi.org/10.1177/1532673X13504414>

Dolan, K. (2014). *When does gender matter? Women candidates and gender stereotypes in American elections*. Oxford University Press.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Flores, A. R. (2015). Attitudes toward transgender rights: Perceived knowledge and secondary interpersonal contact. *Politics, Groups, and Identities*, 3(3), 398–416. <https://doi.org/10.1080/21565503.2015.1050414>

Fulton, S. A. (2012). Running backwards and in high heels: The gendered quality gap and incumbent electoral success. *Political Research Quarterly*, 65(2), 303–314. <https://doi.org/10.1177/1065912911407629>

Haider-Markel, D. P. (2010). *Out and running: Gay and lesbian candidates, elections, and policy representation*. Georgetown University Press.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Huddy, L., & Terkildsen, N. (1993). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science*, 37(1), 119–147.

<https://doi.org/10.2307/2111526>

Kreiss, D., & McGregor, S. C. (2019). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 U.S. presidential cycle. *Political Communication*, 36(2), 155–177.

<https://doi.org/10.1080/10584609.2018.1556350>

Lawless, J. L., & Fox, R. L. (2010). It still takes a candidate: Why women don't run for office. Cambridge University Press.

Mansbridge, J. (2003). Rethinking representation. *American Political Science Review*, 97(4), 515–528.

Mansbridge, J. (1999). Should Blacks represent Blacks and women represent women? A contingent “yes.” *The Journal of Politics*, 61(3), 628–657. <https://doi.org/10.2307/2647821>

McDermott, M. L. (1998). Race and gender cues in low-information elections. *Political Research Quarterly*, 51(4), 895–918. <https://doi.org/10.1177/106591299805100404>

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Pitkin, H. F. (1967). *The concept of representation*. University of California Press.

Reynolds, A. (2013). Representation and rights: The impact of LGBT legislators in comparative perspective. *American Political Science Review*, 107(2), 259–274.

<https://doi.org/10.1017/S0003055413000051>

Sanbonmatsu, K. (2002). Gender stereotypes and vote choice. *American Journal of Political Science*, 46(1), 20–34. <https://doi.org/10.2307/3088412>

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505–523. <https://doi.org/10.1111/rego.12158>

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

APPENDIX A

Experimental Candidate Profiles

The following candidate profiles were presented to each large language model. Across conditions, candidate qualifications, education, occupational experience, endorsements, policy priorities, and campaign messaging were held constant wherever possible. Experimental manipulations altered only identity-related characteristics (e.g., sexual orientation, transgender identity, gender presentation, or leadership style).

Experimental Conditions

Condition	Manipulation
A	Control (non-LGBTQ)
B	Same-sex spouse cue
C	Control
D	Same-sex spouse cue

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

E	Female control
F	Female same-sex spouse cue
G	Female control
H	Female same-sex spouse cue
I	Male control
J	Male same-sex spouse cue
K	Female control
L	Female same-sex spouse cue
M	Explicit transgender identity

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

N	Transgender cue removed
O	Male control
P	Nontraditional leadership + same-sex spouse
Q	Traditional leadership
R	Gender non-conforming manipulation
S	Traditional masculine leadership
T	Nontraditional feminine leadership

APPENDIX C

Outcome Coding

Each AI response produced numerical evaluations.

The following variables were extracted directly from the models:

Algorithmic Electability: Generative AI, Gendered Political Bias, and LGBTQ Candidates

- Electability
- Leadership
- Competence
- Political strength
- Overall candidate quality

When AI systems provided accompanying narrative explanations, those explanations were retained for verification but were not quantitatively coded. Because the primary dependent variables consisted of numerical scores generated directly by the models, intercoder reliability statistics were not applicable.