# From Office to Opportunity: Evaluating the Role of Descriptive Representation on Candidate Emergence in U.S. Swing States*

Jessica Persano†, Jolie Anderson, Xuanting Fan

June 2025

## Abstract

This study investigates whether descriptive representation, specifically defined as the presence of a nonwhite or woman legislator, influences the demographic composition of future candidate pools. Our central theory—the role model effect—proposes that descriptive representatives may inspire others who share their identity to run for office. Using data from 332 state house districts across Georgia, Michigan, and Nevada, we examine whether having a demographically diverse legislator in 2022 is associated with greater racial and gender diversity among 2024 candidates. We estimate a series of multivariate OLS regression models on both the full sample and a subset of open-seat districts where the 2022 legislator did not seek reelection. In this full sample, we find strong evidence that nonwhite and woman legislators are associated with significantly more diverse candidate pools, supporting themes of downstream representational effects. However, these effects largely disappear in open-seat districts, suggesting that observed associations may be driven by incumbents rerunning rather than inspiring new candidates to emerge. Our findings highlight the role of structural factors, such as district demographics, campaign costs, and institutional support, in shaping candidate diversity. We conclude that while descriptive representation can sustain diversity through incumbency, it is not sufficient to broaden access to political office on its own. Expanding diversity in the candidate pipeline will instead require targeted institutional reforms aimed at reducing barriers to entry for underrepresented groups.

**Keywords:** descriptive representation, candidate emergence, political pipeline, political diversity, symbolic politics

---

# 1   Introduction

Descriptive representation refers to the extent to which elected officials physically or demographically resemble the constituents they serve, a concept first articulated by political scientist Hanna Fenichel Pitkin in 1967. Specifically, Fenichel Pitkin argued that descriptive representation occurs when "a person or thing stands for others 'by being sufficiently like them'" (Fenichel-Pitkin, 1967). In this study, we narrowed our focus to two dimensions of identity that can be most easily visualized and understood by constituents: race and gender.

Descriptive representation is widely regarded as essential to a functioning democracy, as it is an important factor in ensuring that elected officials accurately represent the populations that they are governing. Overcoming a history of underrepresentation for minority groups within government positions is an intrinsic part of both ascertaining a diverse legislature and fully achieving representation. Despite steady progress in increasing diverse representation over time, significant disparities remain. As of 2025, only 26% of current U.S. Congress members identify as a race or ethnicity other than non-Hispanic white, even though this group collectively comprises 42% of the U.S. population (Schaeffer, 2025). Similar (if not more extreme) gaps persist at the state level, providing evidence that vast improvements must be made in electing a legislature that reflects the people. Increasing the diversity of the candidate pools is therefore a crucial step towards building more representative legislatures.

This motivation frames the theoretical puzzle of our research. Specifically, we are interested in whether the increased minority voter engagement and political interest that occurs with strong descriptive representation in the state legislature (Rocha et al., 2010) can be extended to candidate emergence. As such, we ask the following research question: Does the presence of demographically diverse state legislators positively influence the demographic diversity of the candidate pool in the subsequent election? We focus on race and gender as key units of analysis, and examine whether the presence of diverse legislators can help 'open the pipeline' for future candidates from underrepresented backgrounds.

This study contributes to a growing body of research showing that descriptive repre-

sentation can shape civic engagement and political outcomes. While prior research has focused primarily on voters, we extend this logic to candidate emergence. Specifically, we test whether having a nonwhite or woman legislator in office increases the likelihood that others from that group will run for office in the next election. Based on the positive impact of descriptive representation on voters, we hypothesize that districts represented by a nonwhite or woman legislator in 2022 will have a higher proportion of nonwhite or women candidates, respectively, in the 2024 elections. Using data from three swing states (Georgia, Michigan, and Nevada), we run a series of ordinary least squares (OLS) regressions to estimate the relationship between legislator demographics in 2022 and candidate pool composition in 2024. Our results show that while having a nonwhite or woman legislator is significantly associated with greater candidate diversity within all districts, these effects diminish or disappear entirely in open-seat districts, suggesting that incumbency plays a key role in reinforcing this pipeline effect.

## 2 Literature Review

In her book, *The Concept of Representation*, Fenichel-Pitkin (1967) introduced the idea of descriptive representation, which occurs when elected officials mirror the demographic characteristics of their constituents. Specifically, Fenichel Pitkin argued that a legislature resembling the broad population, like a map or a mirror might, helps justify and defend the democratic principle of governance by the few over the many (Fenichel-Pitkin, 1967). In this way, descriptive representation serves not only a symbolic role, but also a foundational one in sustaining democratic legitimacy.

However, some scholars have critiqued the large emphasis placed on the importance of descriptive representation, instead highlighting the importance of substantive representation: the alignment between elected officials' policy preferences and those of their constituents (Fenichel-Pitkin, 1967). For example, political scientist Carol Swain, in her 1993 book *Black*

*Faces, Black Interests: The Representation of Black Interests in Congress*, argues that advancing "black policy interests" may depend more on shared policy goals than shared racial identity (Swain, 1993). More recently, Griffin and Newman (2008) argue that descriptive representation does not always translate into substantive responsiveness, particularly when minority legislators face institutional constraints. This debate continues today, as scholars examine the roles that both models of representation play in producing meaningful democratic outcomes.

Nevertheless, a growing body of empirical research has found that descriptive representation can have profound impacts on political behavior and civic engagement. For example, Rocha et al. (2010) demonstrate that African Americans and Latinos were significantly more likely to vote in the 2008 election if they resided in districts where their racial group was descriptively represented compared to districts where it was not, by 20% and 35%, respectively. Other studies have continued to show this positive impact, claiming descriptive representation can lead to high rates of campaign promise fulfillment for women legislators (Homola, 2022) and result in more favorable evaluations of legislators by constituents (Campbell & Cowley, 2014). Taken together, these findings suggest that descriptive representation can generate downstream benefits for both voter participation and political accountability.

However, the relationship between descriptive representation and candidate emergence remains underexplored. The dominant explanations for who runs for office focuses on political pipelines, in which professions like law or business serve as direct stepping stones to elected office - paths that are disproportionately accessible to and dominated by white men (Carnes, 2013). Gendered barriers also persist: women are significantly less likely than men to be encouraged to run for office or to receive recommendations from influential political actors, contributing directly to persistent gender disparities in candidate pools (Lawless & Fox, 2012). Racial barriers seem to mirror this pattern. For example, Ocampo and Ray (2019) found that Latino candidates receive significantly fewer endorsements from parties than non-Latino peers, limiting their ability to emerge as viable candidates.

This literature raises the possibility that descriptive representation might help disrupt these unequal pipelines by serving as a form of political encouragement. One mechanism for this impact is the so-called 'role model effect,' where individuals are more likely to envision themselves running for office when they see someone who looks like them already in that position. Wolbrecht and Campbell (2007) found evidence of this effect in the U.K., where female representation in Parliament was associated with increased political participation and political interest among women and adolescent girls, more so than among male counterparts. This evidence suggests that descriptive representation may not only influence who participates in politics, but also perceptions of political possibility, thus making it a potentially powerful mechanism for diversifying future candidate pools.

Despite these theoretical and empirical foundations, the role of descriptive representation in shaping candidate emergence remains widely underexplored. Prior research has clearly documented the effects of descriptive representation on voter behavior, legislator evaluations, and other forms of civic engagement, but has not adequately addressed whether it encourages individuals from underrepresented groups to run for office themselves. Therefore, our study attempts to fill this gap by examining whether the presence of a demographically diverse legislator increases the demographic diversity of the candidate pool in subsequent elections. In doing so, we attempt to determine whether descriptive representation can play a meaningful role in broadening access to political office and advancing legislative diversity at the state level.

# 3 Theory and Hypotheses

To explain the relationship between legislator identity and subsequent candidate diversity, we propose a *theory of downstream effects of descriptive representation*. Specifically, we theorize that when a legislator shares a key demographic identity, namely race or gender, with underrepresented constituents, it sends a powerful signal of political accessibility. In other

words, seeing someone 'like them' in elected office may encourage individuals from similar backgrounds to view political candidacy as both a viable and attainable path. This symbolic effect is especially relevant for groups historically excluded from political institutions, such as women and racial minorities.

We propose that these downstream effects of descriptive representation may operate through three main mechanisms. First, and most central to our study, is the 'role model effect', or the idea that constituents may feel more motivated to run for office when they see someone who reflects their identity in that position. Second, we theorize that women and minority legislators may be more inclined to support or mentor individuals from similar backgrounds, further expanding access to the political pipeline. Third, political parties and interest groups may interpret the electoral success of diverse legislators as a signal of viability, resulting in the targeted recruitment of more candidates from underrepresented backgrounds. Taken together, these theoretical insights motivate our four primary hypotheses:

**H1:** Districts with a nonwhite legislator in 2022 will have a higher proportion of nonwhite candidates in 2024.

**H2:** Districts with a woman legislator in 2022 will have a higher proportion of women candidates in 2024.

**H3:** These effects will be weaker in districts where the 2022 legislator does not seek re-election (open-seat districts).

**H4:** These effects will remain statistically significant after controlling for other district-level factors, such as racial composition, gender composition, and electoral competitiveness.

# 4   Research Design and Methods

This study evaluates whether descriptive representation influences the demographic composition of future candidate pools by analyzing state house elections in three battleground

states: Georgia, Michigan, and Nevada. These states were selected for their high electoral competitiveness, each having a 2020 presidential margin of victory of less than three percentage points, along with their relatively high levels of citizen diversity. Focusing on these closely contested states allows us to capture candidate dynamics in districts where representation may be especially important for mobilization, recruitment, and competition. The unit of analysis for this project is the state house district. We chose state house districts because their small geographic size and large quantity per state offer a higher level of granularity, allowing for greater variation in candidate dynamics. Furthermore, state house elections occur every two years, making them ideal for observing short-term trends in candidate diversity and the political pipeline.

To measure race and gender diversity in a clear and theoretically meaningful way, we operationalize race as a binary indicator of nonwhite vs. white, and gender as a binary indicator of female vs. male. While this oversimplifies identity and masks within-group variation, it was a necessary choice considering time constraints and data limitations of this project. Nevertheless, this classification more accurately reflects the realities of political visibility; voters and constituents may be more likely to perceive and respond to broad identity categories

As such, the key covariates for our analysis are binary indicators for whether a nonwhite legislator and/or a woman legislator held office in the district after the 2022 election. The outcome variable is the proportion of 2024 candidates in each district who are a) nonwhite and b) women. Confounders that were controlled for include the percent of nonwhite residents and the percent of women residents (from Census ACS estimates) in each district, as well as the district's 2022 percentage point margin of victory, which acts as a proxy for competitiveness. These confounders are included because they are likely to influence both the legislator identity and the demographic makeup of the candidate pool. For example, more diverse districts may be both more likely to elect diverse legislators and attract diverse candidates, while competitive districts may affect recruitment and candidate emergence dy-

namics.

Candidate and legislator data were collected from official election offices and Ballotpedia. For Michigan, legislator data came from the Michigan Legislature (2022) and Ballotpedia (2022b), with candidate data from the Michigan Secretary of State Candidate Listing Database (Office of Secretary of State Jocelyn Benson, 2024) and Ballotpedia (2024b). For Nevada, legislator data came from the Nevada State Legislature (2022) and Ballotpedia (2022c), with candidate data from the Nevada Legislature Research Division (2024) and Ballotpedia (2024c). For Georgia, legislator data came from the Georgia General Assembly (2022) and Ballotpedia (2022a), with candidate data from the Georgia Secretary of State Qualifying Candidate Information Database (Georgia Secretary of State, 2024) and Ballotpedia (2024a). District-level margins of victory were obtained from each state's certified Secretary of State election results (Georgia Secretary of State, 2022; Michigan Secretary of State, 2022; Nevada Secretary of State, 2022). District-level demographic data was drawn from the U.S. Census Bureau's American Community Survey 2021 5-Year Estimates (U.S. Census Bureau, 2022a).

To contextualize our analysis, we provide a summary of the racial and gender composition of candidates and legislators in our dataset. Table 1 displays the percentage of women and nonwhite individuals among all 2024 candidates (807 candidates across 332 districts), non-incumbent 2024 candidates (507 candidates across 256 open-seat districts), and all 2022 state legislators (332 total). Overall, women made up 36.8% of all candidates and 36.5% of non-incumbents, while nonwhite candidates comprised 35.7% and 34.5% of these groups, respectively. Among 2022 legislators, the proportions were slightly higher: 38.9% women and 37.3% nonwhite. These differences suggest that incumbents were modestly more diverse than the broader candidate pool, and that open-seat races may not be expanding access for underrepresented groups. These patterns are explored more formally in the next section, where we use OLS regression to assess whether the presence of a diverse legislator in 2022 predicts candidate diversity in 2024.

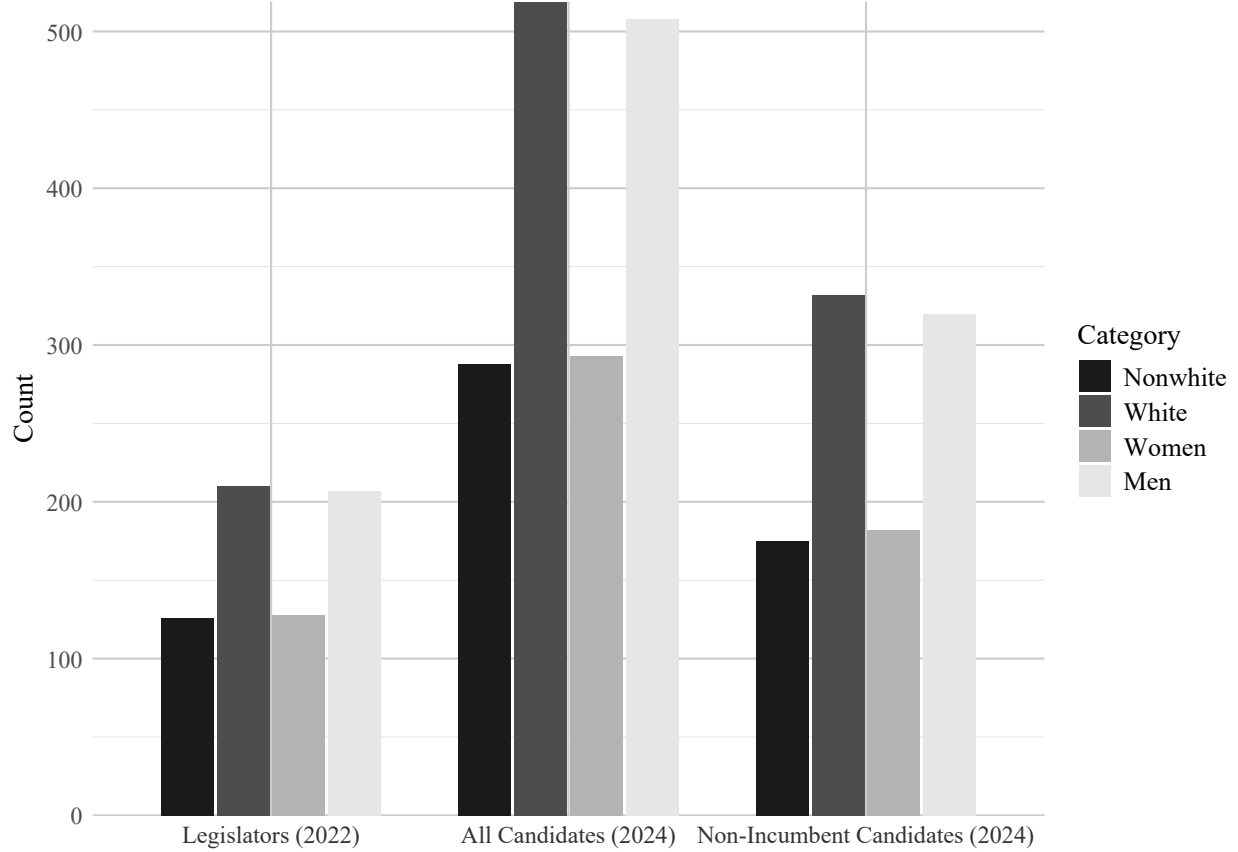Table 1: **Summary of Demographic Composition by Sample**

| Sample | % Women | % Nonwhite | $N$-Value |
|---|---|---|---|
| All Candidates (2024) | 36.8% | 35.7% | 807 |
| Non-Incumbent Candidates (2024) | 36.5% | 34.5% | 507 |
| All Legislators (2022) | 38.9% | 37.3% | 332 |

*Note.* This table summarizes the demographic composition of each sample in the analysis, providing baseline descriptive statistics for assessing trends in candidate diversity among (1) all state house candidates, (2) non-incumbent candidates, and (3) sitting legislators.

Figure 1 visualizes the same data as raw counts, disaggregated by race (nonwhite vs. white) and gender (women vs. men), highlighting the persistent overrepresentation of men and white individuals across all three samples. Taken together, the table and figure provide initial evidence that the underrepresentation of women and nonwhite individuals is not solely a feature of electoral outcomes, but begins with the composition of the candidate pool itself.

Rather than manually coding the race/gender of each legislator and candidate, this study used two algorithmic tools to infer demographic characteristics. Gender was predicted using the Genderize API, which assigns gender probabilities based on first names (Genderize.io, n.d.). Race was predicted using the WRU (Bayesian Improved Surname Geocoding) package, which combines first and last names with geographic information to generate probabilistic estimates of race (Imai & Khanna, 2016). Because WRU performs best with county-level identifiers, each candidate and legislator was assigned to a county based on the geographic overlap between their state legislative district and county boundaries using U.S. Census shapefiles (U.S. Census Bureau, 2022b). Although more precise matching using voter file addresses was considered, it was not feasible within the scope of this project. To assess prediction accuracy, a 10% random sample of names was manually verified using public-facing sources such as campaign websites and news coverage. The results of this validation check are provided in Appendix A. While not perfect, these algorithmic approaches offered a practical solution for generating large-scale identity estimates in a time-efficient and consistent manner.

Figure 1: **Count of Candidates and Legislators by Demographic Group** This bar chart displays the raw count of legislators (2022) and candidates (2024), disaggregated by race (nonwhite vs. white) and gender (women vs. men). Each group is shown for all candidates, non-incumbents only, and legislators.



To analyze our data, we estimate OLS regression with robust standard errors. We run separate models for race and gender. For each outcome, we estimate a bivariate model including only the legislator identity, followed by a multivariate model that includes all three controls. To examine whether effects differ by incumbency, we conduct parallel analyses for (1) all districts and (2) open-seat districts where the 2022 legislator did not seek re-election (non-incumbents). This results in a total of 8 models: bivariate and multivariate regressions for both race and gender, run separately on the full and non-incumbent samples.

# 5 Results

This section presents the results of our empirical analysis assessing whether the presence of a nonwhite legislator and/or woman legislator in 2022 predicts greater demographic diversity in the 2024 candidate pool. For each sample, we first visualize bivariate relationships between legislator identity and candidate diversity, followed by multivariate OLS regression models that control for district demographics and electoral competitiveness. Separate models are estimated for race and gender outcomes, and models are run on two saples: a full sample including all 332 districts, and a sub-sample of only the open-seat districts (N = 256 districts).

We present our main findings below using multivariate OLS regression tables for both the full and non-incumbent district samples. For visual reference, bivariate scatterplots showing the raw relationship between legislator identity and candidate diversity (Figures B1–B4) are included in Appendix B. In the regression tables, we present both bivariate and full models; the coefficient on the legislator identity variable reflects the estimated percentage point change in candidate diversity associated with having a nonwhite or woman legislator in 2022.

## 5.1 Full Sample

We begin with the full sample of 332 districts, which includes all candidates (N = 807 candidates). Regression results for are presented in Tables 2 (race) and 3 (gender), respectively.

As shown in Table 2, having a nonwhite legislator in 2022 is associated with a 58.2 percentage point increase in the share of nonwhite candidates in 2024 in the bivariate model ($\beta = 0.582$, $p < .001$). This relationship remains robust and statistically significant in the full multivariate model ($\beta = 0.427$, $p < .001$), even after accounting for the racial composition of the district and the 2022 margin of victory. The percentage of nonwhite residents is also positively associated with candidate diversity and statistically significant ($\beta = 0.006$, $p < .001$), while margin of victory has no significant effect.

Table 2: **Effect of Nonwhite Legislator on Nonwhite Candidates (Full Sample)**

|  | Nonwhite Candidates (2024) | |
|  | (1) | (2) |
| --- | --- | --- |
| Nonwhite Legislator (2022) | 0.582*** | 0.427*** |
|  | (0.032) | (0.038) |
| % Nonwhite Voters |  | 0.006*** |
|  |  | (0.001) |
| Margin of Victory |  | −0.001 |
| Observations | 332 | 332 |
| R² | 0.500 | 0.563 |
| Adjusted R² | 0.498 | 0.559 |

*p<0.1; **p<0.05; ***p<0.01

*Note.* This table presents OLS regression results predicting the proportion of women candidates in 2024 based on whether the 2022 legislator was a woman, using the full sample of state house districts ($N = 332$). Model 1 is bivariate specification; Model 2 includes controls for district gender composition and margin of victory.

As shown in Table 3 (found on page 12), having a woman legislator in 2022 is associated with a 46.2 percentage point increase in the share of women candidates in 2024 in the bivariate model ($\beta = 0.462$, $p < .001$). This relationship remains robust and statistically significant in the multivariate model ($\beta = 0.454$, $p < .001$), even after accounting for the gender composition of the district and the 2022 margin of victory. The percentage of women residents is also positively associated with candidate diversity ($\beta = 0.018$, $p < .005$), while margin of victory has no significant effect.

## 5.2   Non-Incumbent Sample

Next, we analyze the subset of 256 open-seat districts, where the 2022 legislator did not seek re-election (N = 507 candidates). We analyze open-seat districts separately to assess whether these relationships persist in the absence of the incumbency advantage, where candidacy is open and the role model effects we propose could seemingly have the largest impact. Regression results are presented in Tables 4 (race) and 5 (gender), respectively.

Table 3: **Effect of Woman Legislator on Women Candidates (Full Sample)**

|  | Women Candidates (2024) | |
|  | (1) | (2) |
| --- | --- | --- |
| Woman Legislator (2022) | 0.462*** | 0.454*** |
|  | (0.031) | (0.031) |
| % Women Voters |  | 0.018** |
|  |  | (0.009) |
| Margin of Victory |  | −0.0001 |
|  |  | (0.0004) |
| Observations | 332 | 332 |
| R$^2$ | 0.409 | 0.416 |
| Adjusted R$^2$ | 0.407 | 0.411 |

$^*$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Note.* This table presents OLS regression results predicting the proportion of women candidates in 2024 based on whether the 2022 legislator was a woman, using the full sample of state house districts ($N = 332$). Model 1 is a bivariate specification; Model 2 includes controls for district gender composition and margin of victory.

As shown in Table 4 (found on page 13), the relationship between having a nonwhite legislator in 2022 and the share of nonwhite candidates in 2024 weakens substantially in when only analyzing the sample of open-seat districts (N = 256). The bivariate coefficient decreases to 0.276 ($p < .001$), and becomes statistically insignificant in the full model ($\beta = 0.045$, $p > .05$). The percentage of nonwhite residents remains a significant yet substantively limited predictor of candidate diversity ($\beta = 0.009$, $p < .001$), while margin of victory has no statistically significant effect.

As shown in Table 5 (found on page 14), there is similarly no statistically significant association between having a woman legislator in 2022 and the share of women candidates in open-seat districts, once again showing a significantly weaker relationship than the model including all districts. The coefficient is negative and close to zero in both the bivariate ($\beta = -0.008$, $p > .05$) and multivariate ($\beta = -0.006$, $p > .05$) models. Neither the percentage of women residents nor margin of victory significantly predicts candidate diversity in this subset.

Table 4: **Effect of Nonwhite Legislator on Nonwhite Candidates (Non-Incumbent)**

|  | Nonwhite Candidates (2024) | |
| --- | --- | --- |
|  | (1) | (2) |
| Nonwhite Legislator (2022) | 0.276*** | 0.045 |
|  | (0.053) | (0.061) |
| % Nonwhite Voters |  | 0.009*** |
|  |  | (0.001) |
| Margin of Victory |  | −0.001 |
|  |  | (0.001) |
| Observations | 256 | 256 |
| $R^2$ | 0.098 | 0.233 |
| Adjusted $R^2$ | 0.094 | 0.223 |

$^*$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Note.* This table presents OLS regression results predicting the proportion of non-white candidates in 2024 based on whether the 2022 legislator was nonwhite, using the sample of only open-seat state house districts ($N = 256$). Model 1 is a bivariate specification; Model 2 includes controls for district racial composition and margin of victory.

Overall, the results indicate that the presence of demographically diverse legislators is significantly, though modestly, associated with greater diversity in the subsequent candidate pool. However, this relationship weakens substantially, and in some cases disappears entirely, in open-seat districts. This pattern underscores the important role incumbency may play in reinforcing and sustaining pipelines of diverse political representation. A fuller exploration of these dynamics is provided in the discussion section.

# 6   Discussion

The results of this study offer mixed support for the theory that descriptive representation fosters more diverse candidate pipelines. While we find clear evidence that having a nonwhite or woman legislator in 2022 is associated with greater candidate diversity in 2024, this relationship weakens substantially, or disappears entirely, once incumbents are removed from the analysis. These patterns offer important insights into the mechanisms through which

Table 5: **Effect of Woman Legislator on Women Candidates (Non-Incumbent)**

|  | Women Candidates (2024) | |
| --- | --- | --- |
|  | (1) | (2) |
| Woman Legislator (2022) | −0.008 | −0.006 |
|  | (0.052) | (0.052) |
| % Women Voters |  | −0.012 |
|  |  | (0.016) |
| Margin of Victory |  | 0.0004 |
|  |  | (0.001) |
| Observations | 256 | 256 |
| R$^2$ | 0.0001 | 0.003 |
| Adjusted R$^2$ | -0.004 | -0.009 |

$^*$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

*Note.* This table presents OLS regression results predicting the proportion of women candidates in 2024 based on whether the 2022 legislator was a woman, using the sample of only open-seat state house districts ($N = 256$). Model 1 is a bivariate specification; Model 2 includes controls for district gender composition and margin of victory.

descriptive representation operates, and helps clarify some of the limits of proposed symbolic and inspirational effects on candidate emergence - raising important implications for both democratic theory and political recruitment practices.

In the full sample, the presence of a nonwhite or woman legislator significantly predicted a higher proportion of candidates from that group in the following election. Specifically, having a nonwhite legislator was associated with a 42.7 percentage point increase in the share of nonwhite candidates, and having a woman legislator was associated with a 45.4 percentage point increase in the share of women candidates. These large coefficients likely reflect the small candidate pools per district (average of 1-3 candidates per district), meaning that one additional diverse candidate has an outsized effect on overall proportions. Nevertheless, these findings are consistent with our primary hypotheses and extend prior research by showing that descriptive representation may shape not only voter behavior but also candidate dynamics. However, the most plausible interpretation of these results is not necessarily that diverse legislators inspired new candidates to run, but rather that they chose to run

again themselves. In other words, the strong effects in the full sample are likely driven by incumbents running for office, thereby contributing directly to increased candidate diversity.

This interpretation is reinforced by the results in the non-incumbent sample. Once incumbent legislators are removed from the candidate pool, the relationship between legislator identity and candidate diversity weakens dramatically (in the case of race) and disappears entirely (in the case of gender). In the multivariate models, neither identity variable remains statistically significant, suggesting that other district-level factors may explain most of the variation. In fact, the percent of nonwhite residents is a significant and positive predictor of candidate emergence, implying that the demographic composition of the district itself, rather than the identity of the legislator, better predicts candidate diversity in open-seat districts. In these contexts, diverse communities may organically generate diverse candidate pools, particularly for racial minorities, even in the absence of a racially representative incumbent. Interestingly, margin of victory had no clear effect in any model, suggesting that competitiveness alone may not shape the emergence of diverse candidates once other factors are accounted for.

Importantly, this shift may also reflect the reduced sample size of the non-incumbent sample. Open-seat contests accounted for only 256 of the 332 districts in our sample, reducing the precision of coefficient estimates. However, the dramatic decline in effect size suggests that this is not only an issue of statistical power. Rather, it underscores that the presence of a nonwhite or woman legislator does not appear to inspire new candidates from that same group once that legislator leaves office. Theoretically, this suggests that studies of representation should carefully distinguish between the direct effects of incumbency and broader symbolic influence, as conflating the two may overstate the true impact of descriptive representation.

Still, even if the results are driven by incumbents re-running rather than by new candidates entering the pool, our findings remain normatively important. In particular, they suggest that diverse incumbent legislators may feel sufficiently empowered, supported, or

electorally viable to pursue another term, conditions that are themselves critical to sustaining representation. Additionally, the continued presence of diverse incumbents in office allows constituents to keep benefitting from the downstream effects of descriptive representation on trust, engagement, and accountability.

Another notable pattern is that the effects of legislator race were slightly stronger than those for gender across all models, which may reflect structural differences in how racial vs. gender identity functions in the political arena. Although speculative, we theorize that racial identity could operate as a more visible and salient marker of group membership within political contexts. Racial minority candidates may face greater barriers to entry, but once they are in office, they may serve as especially powerful symbols of inclusion. Gender, by contrast, is more evenly distributed across districts and thus may carry less symbolic weight in certain contexts. Additionally, gendered expectations about leadership and electability of women may limit the extent to which women legislators inspire other women to run, even when they hold office.

Together, these findings suggest that the role model effect, while compelling in theory, may be more limited in practice. Descriptive representation does not automatically translate into broader diversity in the candidate pool. Instead, it appears to reinforce itself through incumbency, sustaining representation when diverse legislators run again but failing to generate new candidacies from underrepresented groups in their absence. This pattern points to the strong influence of structural barriers, such as resource disparities, recruitment gaps, and campaign costs, that pose major barriers for new diverse candidates to enter the political arena.

In sum, our results provide modest support for a theory of downstream descriptive effects, but only in contexts where diverse legislators run again. The absence of effects in open-seat contests highlight the limitations of relying solely on descriptive representation to achieve broader candidate diversity. Practically, if descriptive representation does not open pathways for new candidates, it may fail to fulfill its democratic promise of equal access

to political power. Thus, to build more inclusive political pipelines and achieve more equitable representation, political institutions must pair symbolic representation with targeted structural interventions - such as party mentorship programs, funding support for first-time candidates, and deliberate recruitment of women and candidates of colors. Descriptive representation can help sustain progress, but it cannot, on its own, open the door for the next generation.

# 7  Conclusion

This study set out to investigate whether descriptive representation, specifically the presence of a nonwhite or woman legislator, leads to more diverse candidate pools in the next election cycle. Overall, the findings provide qualified support for this theory, but with important caveats. In the full sample of districts, we found strong evidence that having a nonwhite or woman legislator in 2022 was associated with a significantly higher proportion of nonwhite or women candidates in 2024, aligning with the theory of downstream effects of descriptive representation described above.

However, these effects largely disappeared when we examined open-seat districts alone, with the relationship between legislator diversity and candidate diversity becoming statistically insignificant in the multivariate models. This suggests that our observed relationships may be driven primarily by the decision of incumbents to run again, rather than a broader shift in who feels empowered to run. While this was not the inspirational shift we theorized, it still remains normatively important: if incumbents are re-running, this may signal institutional support for and long-term retention of underrepresented legislators.

Notably, the race models also produced slightly stronger effects than gender models, which may reflect the heightened visibility and salience of racial identity compared to gender identity in political settings. We also found that district-level racial composition was a significant predictor of candidate diversity for non-incumbent samples, indicating that envi-

17

ronmental factors may also play a strong role alongside symbolic ones in shaping who decides to run.

These results have practical implications for parties, advocacy groups, and policymakers aiming to diversify political pipelines. Supporting diverse incumbents through methods such as mentorship, funding, and institutional backing, may be the most reliable way to preserve candidate diversity over time. However, expanding the candidate pipeline will require additional structure reforms, such as candidate training programs, targeted campaign funding, and policies that reduce barriers to entry for first-time candidates from underrepresented groups.

Like all studies, this analysis comes with several limitations. First, the study is observational and cross-sectional, meaning the results are correlational rather than causal; we cannot definitively conclude that descriptive representation caused changes in candidate diversity. Second, despite controlling for some district-level characteristics, we could not account for other key confounders, such as party recruitment dynamics, campaign costs, or access to mentorship, that directly influences who runs. Third, our sample is limited to three swing states, which restricts the generalizability of our findings to broader political contexts. Fourth, our race and gender variables rely on algorithmic predictions from names and geography, and are limited to binary categories. These simplifications prevent deeper analyses of intra-group variation and intersectional identities, and may lead to misclassification in some cases. Finally, the relatively small sample of districts, particularly in the non-incumbent models, limits statistical power and may obscure subtle but meaningful patterns in candidate emergence.

Therefore, future research should build on this work by expanding the dataset to include more states and more election cycles, ideally incorporating longitudinal designs that track candidate emergence over time to allow for a more dynamic and comprehensive understanding of descriptive representation, legislator retention, and candidate emergence. Future research should also work to collect more precise identity data and explore intersectional pat-

terns of candidacy that move beyond binary categories. If possible, future research may also incorporate experimental or quasi-experimental designs to attempt to identify causal pathways, such as matched comparisons of districts. Finally, future research should continue investigating structural barriers and recruitment mechanisms, such as party gatekeeping or campaign financing, that may influence both who chooses to run and who receives support.

In sum, our findings suggest that descriptive representation matters, but its effects are widely conditional. While it may help sustain diversity through incumbency, it is not a substitute for broader reforms to make political candidacy more accessible. Instead, meaningful gains in diverse representation will require not only symbolic breakthroughs, but sustained institutional commitments to increasing equity in the recruitment and retention of under-represented leaders.

# References

Ballotpedia. (2022a). *Georgia House of Representatives Elections, 2022*. Retrieved June 10, 2025, from https://ballotpedia.org/Georgia_House_of_Representatives_elections,_2022

Ballotpedia. (2022b). *Michigan House of Representatives Elections, 2022*. Retrieved June 10, 2025, from https://ballotpedia.org/Michigan_House_of_Representatives_elections,_2022

Ballotpedia. (2022c). *Nevada State Assembly Elections, 2022*. Retrieved June 10, 2025, from https://ballotpedia.org/Nevada_State_Assembly_elections,_2022

Ballotpedia. (2024a). *Georgia House of Representatives Elections, 2024*. Retrieved June 10, 2025, from https://ballotpedia.org/Georgia_House_of_Representatives_elections,_2024#Primary

Ballotpedia. (2024b). *Michigan House of Representatives Elections, 2024*. Retrieved June 10, 2025, from https://ballotpedia.org/Michigan_House_of_Representatives_elections,_2024#Primary

Ballotpedia. (2024c). *Nevada State Assembly Elections, 2024*. Retrieved June 10, 2025, from https://ballotpedia.org/Nevada_State_Assembly_elections,_2024#Primary

Campbell, R., & Cowley, P. (2014). What Voters Want: Reactions to Candidate Characteristics in a Survey Experiment. *Political Studies*, *62*(4), 745–765. https://doi.org/10.1111/1467-9248.12048

Carnes, N. (2013). *White-Collar Government: The Hidden Role of Class in Economic Policy Making*. University of Chicago Press.

Fenichel-Pitkin, H. (1967). *The Concept of Representation*. University of California Press.

Genderize.io. (n.d.). *Genderize API*. Demografix ApS. https://genderize.io

Georgia General Assembly. (2022). *House Members*. Retrieved June 10, 2025, from https://www.legis.ga.gov/members/house?session=1031

Georgia Secretary of State. (2022). *2022 General Election Results*. Retrieved June 10, 2025, from https://results.sos.ga.gov/results/public/Georgia/elections/2022NovGen

Georgia Secretary of State. (2024). *Qualifying Candidate Information Database*. Retrieved June 10, 2025, from https://mvp.sos.ga.gov/s/qualifying-candidate-information

Griffin, J. D., & Newman, B. (2008). *Minority Report: Evaluating Political Equality in America*. University of Chicago Press.

Homola, J. (2022). The Effects of Women's Descriptive Representation on Government Behavior. *Legislative Studies Quarterly*, *47*(2), 295–308. https://doi.org/10.1111/lsq.12330

Imai, K., & Khanna, K. (2016). Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records. *Political Analysis*, *24*(2), 263–272. https://doi.org/10.1093/pan/mpw001

Lawless, J., & Fox, R. L. (2012). *Men Rule: The Continued Under-Representation of Women in U.S. Politics*. Women & Politics Institute, American University. https://www.american.edu/spa/wpi/upload/2012-men-rule-report-final-web.pdf

Michigan Legislature. (2022). *House of Representatives: 2023–2024 Session*. Retrieved June 10, 2025, from https://www.legislature.mi.gov/Legislature/Legislators?session=2023-2024

Michigan Secretary of State. (2022). *Election Results and Vote History*. Retrieved June 10, 2025, from https://mvic.sos.state.mi.us/votehistory/Index?type=C&electionDate=11-5-2024/

Nevada Legislature Research Division. (2024). *2024 Primary Election Nevada Legislative Candidates*. Retrieved June 10, 2025, from https://www.leg.state.nv.us/Division/Research/Content/items/2024-primary-election-nevada-legislative-candidates

Nevada Secretary of State. (2022). *Assembly Election Results*. Retrieved June 10, 2025, from https://silverstateelection.nv.gov/NVAssembly/

Nevada State Legislature. (2022). *Assembly Members* [archived web page]. Internet Archive Wayback Machine. Retrieved June 10, 2025, from https://web.archive.org/web/20230112053652/https://www.leg.state.nv.us/App/Legislator/A/Assembly/Current

Ocampo, A. X., & Ray, J. (2019). Many Are Called But Few Are Chosen: The Emergence of Latino Congressional Candidates. *Politics, Groups, and Identities, 8*(4), 738–761. https://doi.org/10.1080/21565503.2019.1629968

Office of Secretary of State Jocelyn Benson. (2024). *2024 Primary Candidate Listing Database.* Retrieved June 10, 2025, from https://mi-boe.entellitrak.com/candlist/2024PRI_CANDLIST.html

Rocha, R. R., Tolbert, C. J., Bowen, D. C., & Clark, C. J. (2010). Race and Turnout: Does Descriptive Representation in State Legislatures Increase Minority Voting? *Political Research Quarterly, 63*(4), 890–907. https://doi.org/10.1177/1065912910376388

Schaeffer, K. (2025, January 21). *119th Congress Brings New Growth in Racial, Ethnic Diversity to Capitol Hill.* Pew Research Center. Retrieved July 25, 2025, from https://www.pewresearch.org/short-reads/2025/01/21/119th-congress-brings-new-growth-in-racial-ethnic-diversity-to-capitol-hill/

Swain, C. M. (1993). *Black Faces, Black Interests: The Representation of African Americans in Congress.* Harvard University Press.

U.S. Census Bureau. (2022a). *American Community Survey 5-Year Estimates, 2017–2021* [dataset]. https://www.census.gov/programs-surveys/acs

U.S. Census Bureau. (2022b). *TIGER/Line Shapefiles* [dataset]. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

Wolbrecht, C., & Campbell, D. E. (2007). Leading by Example: Female Members of Parliament as Political Role Models. *American Journal of Political Science, 51*(4), 921–939. https://doi.org/10.1111/j.1540-5907.2007.00289.x

# A    Appendix A: Manual Race & Gender Algorithm Check

This appendix provides an overview of a manual validation check on the race/gender algorithms. To assess the accuracy of our demographic prediction methods, we conducted a manual validation of both the Genderize API (used for gender prediction) and the WRU package (used for race prediction). A random 10% sample of legislators (n = 30) and candidates (n = 75) were drawn from our full dataset. The validation sample was stratified across all three states, and cases were randomly assigned to reviewers. Reviewers confirmed whether the predicted gender and race matched publicly identifiable characteristics using publicly available information, such as campaign websites, official biographies, photos, and news coverage. Verification was cross-checked by multiple team members when needed. In cases where the algorithm was incorrect, the correct classification was noted.

Gender prediction accuracy was 100% for legislators and 96% for candidates, with only three misclassifications (all due to unusual or ambiguous names such as Mishon or Adeline). Race prediction accuracy was 83.3% for legislators and 84% for candidates. Most errors involved individuals misclassified as Black when they were white, or vice versa. Notably, race misclassification errors were most frequent among legislators and candidates in Georgia. While this partly reflects Georgia's larger share of the validation sample, error rates were also highest in Georgia, suggesting the WRU algorithm may have performed less reliably in this state's naming and racial context. One possible explanation for this is that local naming patterns in Georgia, particularly among Black legislators/candidates, may diverge from national surname-race correlations used by the WRU algorithm. Additionally, county-level racial demographics in Georgia may be less predictive of individual race due to high levels of racial heterogeneity within counties along with possible district-county mismatch.

While the overall accuracy rates were relatively high, especially for gender, these findings underscore the need for caution in interpreting results. The binary coding system and

probabilistic nature of WRU can obscure within-group diversity and may introduce minor biases. However, given the scale of the data and resource constraints, these algorithms remain a helpful method for approximating descriptive characteristics in this type of research.

# B   Appendix B: Bivariate Scatterplots of 2022 Legislator and 2024 Candidate Diversity

This appendix presents the bivariate relationships between race/gender of 2022 state legislators and the proportion of nonwhite/women candidates running in 2024. Each scatterplot visualizes data at the district level, with jitter applied to binary x-axis values to reduce overlap and 95% confidence intervals included to show the direction and strength of the relationship. These visualizations supplement the multivariate regression results presented in the main body of the paper. To ensure readability, each plot is placed on its own page.

Figure A1: **Effect of Nonwhite Legislator on Nonwhite Candidates (Full Sample)**
This scatterplot displays the relationship between the race of the 2022 legislator and the share of nonwhite candidates running in 2024 across all districts (N = 332). A regression line with 95% confidence intervals is included. Jitter is used to minimize point overlap.
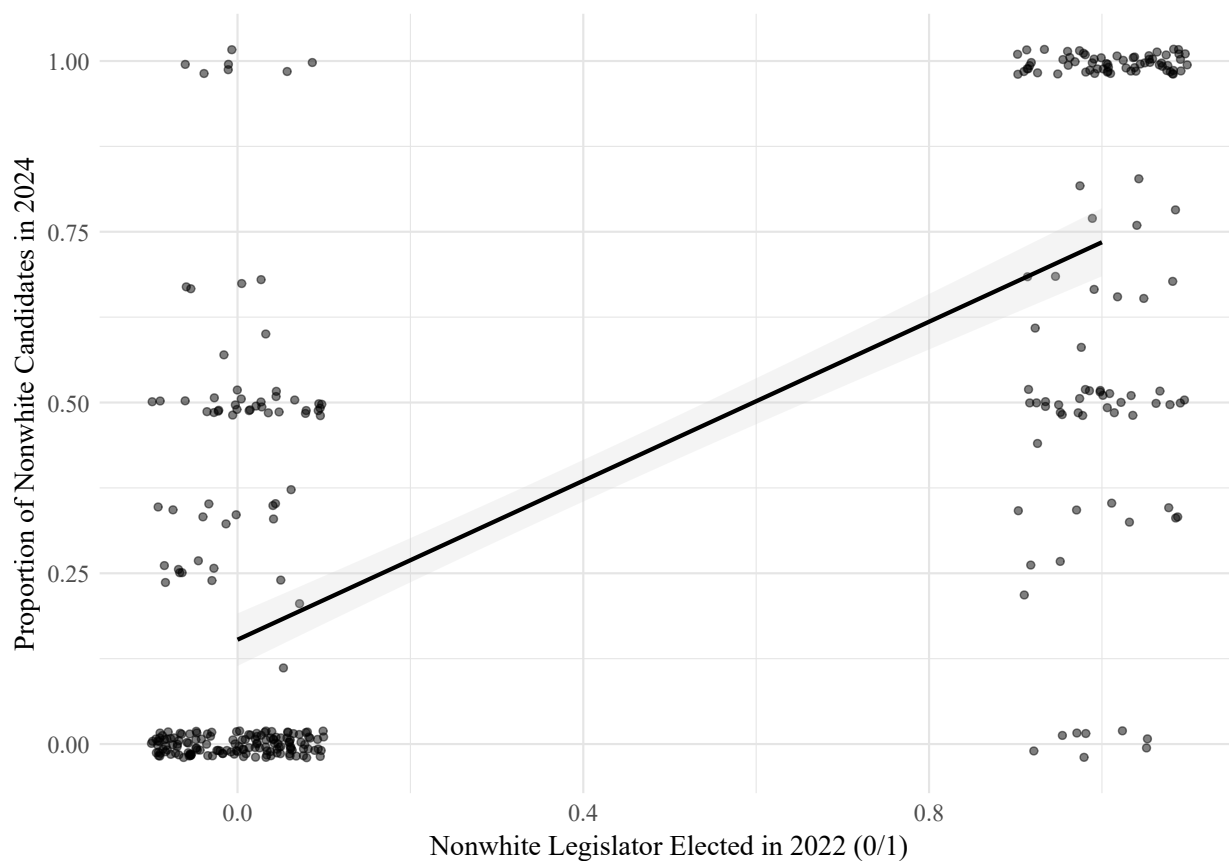
Figure A2: **Effect of Woman Legislator on Women Candidates (Full Sample)** This scatterplot displays the relationship between the gender of the 2022 legislator and the share of women candidates running in 2024 across all districts (N = 332). A regression line with 95% confidence intervals is included. Jitter is used to minimize point overlap.
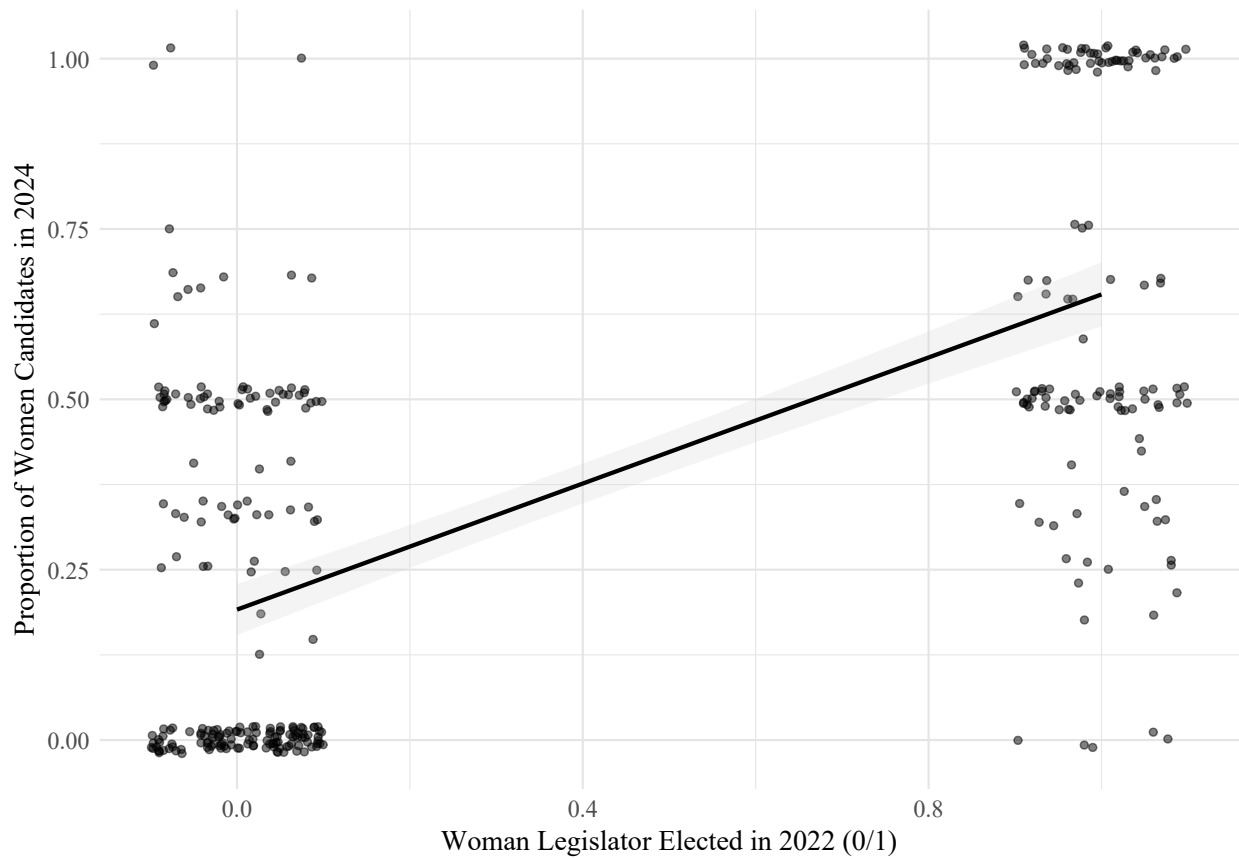
Figure A3: **Effect of Nonwhite Legislator on Nonwhite Candidates (Non-Incumbent Sample)** This scatterplot displays the relationship between the race of the 2022 legislator and the share of nonwhite candidates running in 2024 among open-seat districts (N = 256). A regression line with 95% confidence intervals is included. Jitter is used to minimize point overlap.
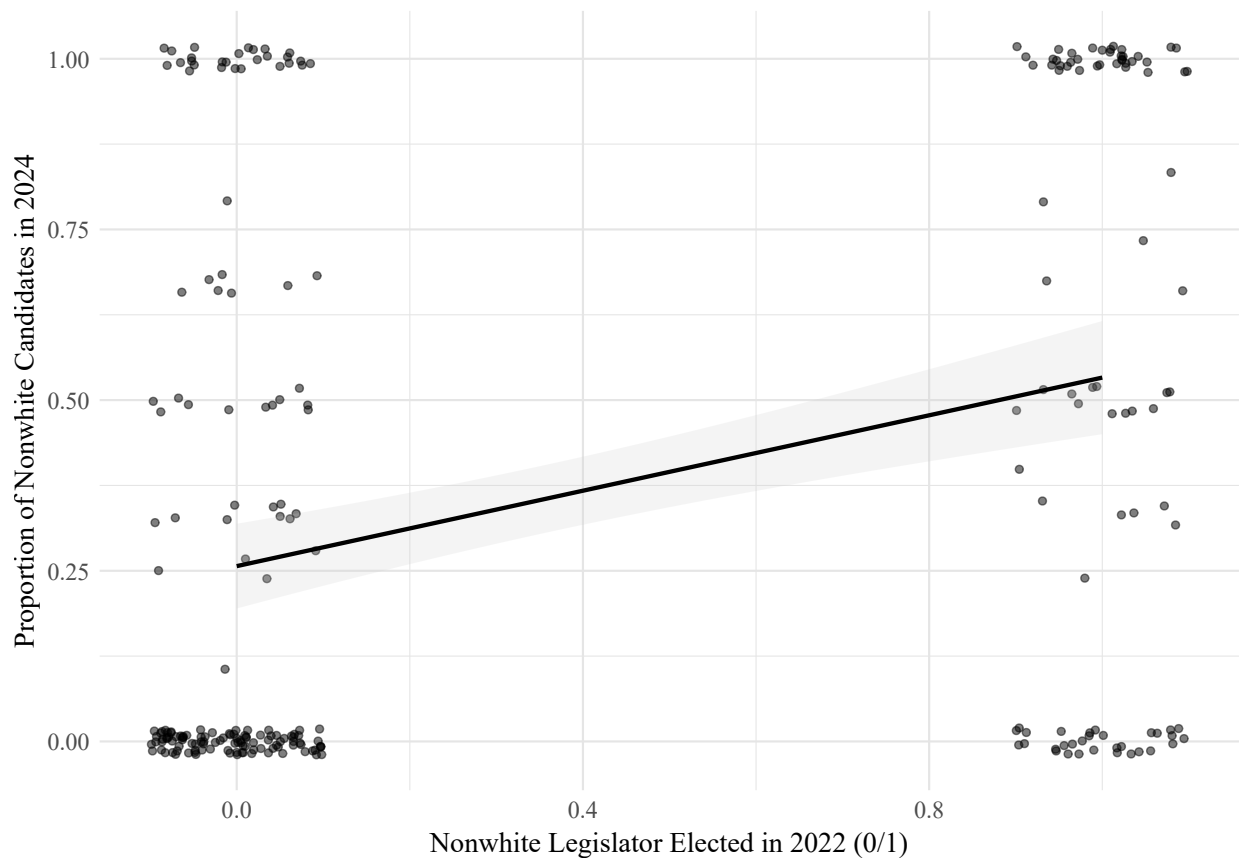
Figure A4: **Effect of Woman Legislator on Women Candidates (Non-Incumbent Sample)** This scatterplot displays the relationship between the gender of the 2022 legislator and the share of women candidates running in 2024 among open-seat districts (N = 256). A regression line with 95% confidence intervals is included. Jitter is used to minimize point overlap.