Racial Self-Classification, Group Consciousness, and Public Employment Representation

Diogo Baerlocher
University of South Florida

Rodrigo Schneider
Skidmore College

September, 2025*

Abstract

This paper examines how racial identity misrepresentation influences public sector hiring in Brazil. We focus on misaligned white candidates — those who self-identify as white but are unlikely to be classified as such by facial recognition — and exploit close electoral races using a regression discontinuity design. Narrow victories by these candidates reduce the share of nonwhite hires in municipal legislative offices by approximately 20%, with effects concentrated in temporary and managerial positions. We also find a significant decline in nonwhite leadership in municipal secretariats. These results indicate that misaligned whiteness shapes racial representation through political and bureaucratic channels.

Keywords: Racial classification, Political representation, Phenotypic discrimination, Public employment

Baerlocher: Department of Economics, University of South Florida, 4202 E. Fowler Ave. CMC342, Tampa, FL, 33620. (email: baerlocher@usf.edu). *Schneider:* Department of Economics, Skidmore College, 223 Filene Hall, 815 N Broadway, Saratoga Springs, NY, 12866 USA. (email: rschnei2@skidmore.edu).

1 Introduction

Racial inequality remains a persistent feature of many democracies, and political representation is often viewed as a critical avenue for addressing it. The race and ethnicity of elected officials are thought to influence political behavior, policy outcomes, and the extent to which historically marginalized groups feel included in the political process. However, racial self-identification is not always straightforward—particularly in societies whose colonial histories led to widespread racial mixing and the formation of social hierarchies that privilege dominant racial groups. In such contexts, racial self-classification and racial group consciousness may diverge, potentially limiting the transformative potential of descriptive representation.

In this paper, we examine how the absence of racial group consciousness among elected city councilors relates to policy outcomes in the Brazilian context. We define city councilors as lacking racial group consciousness when they self-identify as *white* but possess phenotypical characteristics that make it unlikely for them to be perceived as such. Throughout the paper, we refer to these individuals as *misaligned white*. To assess phenotypical appearance, we apply a facial recognition algorithm to candidate photographs, which assigns probabilities of classification into six ethnoracial categories. Using these outputs, we calculate each candidate's predicted probability of being perceived as *white* and identify those who self-reported as *white* despite having a low predicted probability of being perceived that way. Focusing on the 2020 municipal elections for city councils, we find that municipalities where *misaligned white* candidates were elected by a narrow margin have a lower share of *nonwhite* public employees, suggesting that the absence of racial group consciousness may limit racial representation in public employment.

Our empirical strategy relies on a Regression Discontinuity Design, comparing municipalities where *misaligned white* candidates were narrowly elected to the city council with those where *misaligned white* candidates narrowly lost. We show that, near the cutoff, the individual characteristics of the candidates — as well as the characteristics of the municipalities — are similar. The only exception is population size, which is larger in municipalities where *misaligned white* candidates were elected. However, we demonstrate that our results are robust to controlling for this variable. Data on the margin of victory, self-reported race, and other socioeconomic characteristics of the candidates come from the Brazilian Superior Electoral Court, while our primary source of data on public employment is the Annual Report of Social Information, maintained by Brazil's Ministry of Economy.

We also present an additional set of analyses to strengthen the validity and interpretation of our findings.

We show that the effect of electing a *misaligned white* politician is particularly pronounced in the racial composition of temporary public workers, rather than permanent workers, who possess tenure and are more difficult to replace. This result reinforces the interpretation that the observed effects stem from political actions undertaken by *misaligned white* politicians, rather than from structural features that our empirical strategy could not fully account for. Moreover, we find that the effect holds for temporary workers in both managerial and non-managerial positions. Finally, we examine changes in leadership by comparing the racial composition of the heads of the health and education secretariats before and after the 2020 election. We find that electing one additional *misaligned white* politician reduces the probability that a municipality experiences an increase in the number of *nonwhite* managers in the health secretariat, but we find no significant effect for the education secretariat. However, the estimated effects on the heads of the secretariats are not robust to alternative face-detection algorithms and should be interpreted with caution.

A key feature of the Brazilian context is the fluidity of racial classification. Since 2014, political candidates have been required to report their race, yet self-identification often diverges from social perception. For example, Janusz (2022) finds that roughly one-quarter of politicians in their sample reported a racial group different from the one perceived by a panel of coders. Further, Janusz (2021) documents that 27% of politicians who ran in both 2014 and 2016 changed their self-reported race, and Romero (2021) reports that over 4,000 candidates reclassified themselves from *white* to *black* in the 2020 elections. These patterns highlight how, in Brazil, race operates as a social construct (Sen and Wasow, 2016).

Brazil's understanding of race is deeply rooted in its unique historical formation. Portuguese colonization, African slavery, and the presence of Indigenous peoples produced a broad range of phenotypical characteristics, which form the main basis for racial classification in Brazil.[1] After the abolition of slavery in the late nineteenth century, the state pursued a "whitening" policy, notably removing the *pardo* category from the 1890 Census and classifying mixed-race individuals as *white* to inflate the proportion of *whites* (Camargo, 2009). Historically, this contributed to mixed-race Brazilians' tendency to self-identify as *white*.[2] From the 1990s onward, Afro-Brazilian organizations advanced self-identification based on ethnoracial identity, and affirmative action policies reinforced this movement. As a result, the 2022 Census, for the first time, reported a majority identifying as *pardo* (Pinhoni and Croquer, 2023). This trajectory underpins our link between

[1] Unlike in the U.S., where race is largely determined by ancestry (Skidmore, 2009, p. 5).

[2] This tendency helped prompt the removal of race-related questions from several censuses. As Camargo (2009) notes, the 1920 Census omitted a race question altogether, while the 1940 Census replaced it with a "color" question and excluded *pardo*. Since 1950—except in 1970—the census has included color (and "color and race" since 1991) with *white*, *black*, *yellow*, and *brown* (*pardo*) as main categories.

racial group consciousness and the statistical definition of *misaligned white* candidates, who likely self-report based on historical and social constructs rather than phenotypical appearance, distancing themselves from recent Afro-Brazilian advocacy efforts.

Related Literature Our paper contributes to the literature on the effects of ethnoracial representation on constituents. For example, Beach and Jones (2017) show that greater ethnic diversity in California city councils leads to reduced public spending, likely due to increased disagreement among legislators—consistent with the classic findings of Alesina, Baqir and Easterly (1999). In a related study, Beach, Jones, Twinam and Walsh (2024) find that the election of *nonwhite* city council members led to larger gains in housing values in *nonwhite* neighborhoods compared to *white* ones. Focusing on labor markets in large U.S. cities, Nye, Rainer and Stratmann (2015) show that *black* employment and labor force participation increase following the election of *black* mayors.

In the Brazilian context, de Lucena Coelho, Estevan, Nakaguma and Rabelo (2024) find that the election of *nonwhite* mayors does not significantly affect the employment of *nonwhite* individuals in top government positions, whereas Ikawa, Martins, Sant'Anna and Santarrosa (2024) shows that the election of *black* mayors leads to higher enrollment and graduation rates among *black* students. The study by de Lucena Coelho et al. (2024) is particularly relevant to ours, as it also examines the impact of local elections on public employment and relies on candidate photographs. However, unlike our approach, they use photographs to infer candidates' race in the absence of self-reported data, while we use photographs to identify inconsistencies between self-reported race and phenotypical characteristics. Another key difference is that we focus on the election of city council legislators, whereas both cited studies examine mayoral elections.

We focus on city councilors rather than mayors for two main reasons. First, recent research highlights the pivotal role of councilors in shaping local governance. They exert direct influence over temporary hiring within their jurisdictions and indirect influence over executive appointments—such as municipal secretariats — by leveraging their control over budget approval and coalition support to secure politically valuable contracts. For instance, Colonnelli, Prem and Teso (2020) shows that discretionary public employment in Brazil serves as a patronage resource. Other studies demonstrate that councils shape fiscal outcomes through their authority over taxation and spending (Schneider and Veras, 2023), and can even affect homicide rates, despite security being a state-level responsibility (Novaes, 2024). Second, our analysis of councilors is motivated by sample size considerations. Because information on the race of public employees is only available for 2023, we must link employment outcomes to the 2020 elections. This makes it difficult to rely solely on mayors, whose numbers are limited once restrictions related to phenotypical appearance are applied. Focusing on councilors therefore provides a much larger and more informative sample.

Finally, our paper is related to the concept of racial group consciousness studied in political science. In particular, we adopt the definition used by Smith, Clemons, Krishnamurthy, Martinez, McLaren and White (2024), in which racial group consciousness is an individual's "realization of their position as a disadvantaged racial group and their willingness to use politics to improve this position" (p. 1672). While prior research has focused on how such consciousness shapes voter behavior — often with reference to the American Civil Rights Movement (*e.g*, Cascio and Washington, 2014) — we shift the lens to politicians themselves. In our framework, *misaligned white* politicians lack this awareness, as they do not see themselves as part of a disadvantaged racial group.

2 Empirical Strategy

Our objective is to estimate the effects of electing municipal councilors who lack racial group consciousness on the racial composition of municipal hiring. We define such a candidate as one who self-reports as *white* but whose phenotypical appearance suggests otherwise. We refer to these candidates as *misaligned whites* and describe how we identify them in detail below. This definition rests on the assumption that *misaligned white* candidates report their race in line with historical whitening tendencies, rather than in response to recent movements promoting racial awareness.

Our empirical strategy is a cross-sectional regression discontinuity (RD) design:

$$Y_m = \alpha + \lambda D_m + \beta_1 M_m + \beta_2 (M_m \times D_m) + \epsilon_m, \tag{1}$$

where Y_m denotes the racial composition of municipal hires in municipality m, D_m is a binary indicator equal to one if a *misaligned white* candidate was elected, and $M_m \in [-h, h]$ is the margin of victory centered at zero. The error term is ϵ_m, and the bandwidth h is selected optimally. If multiple *misaligned white* candidates run in the same municipality, we use the margin of the candidate closest to the cutoff.

Data. Our primary source for data on the racial composition of municipal hiring is the Annual Report of Social Information (RAIS), maintained by Brazil's Ministry of Economy. The RAIS contains detailed information on formal-sector employees, including wages, hours worked, occupation, sector of employment, and race. While most information is reported by the employer, sociodemographic fields such as race are provided by the worker at the time of hiring. Further details on RAIS collection procedures are discussed in Cornwell, Rivera and Schmutte (2017). We focus on the 2023 data — the third year of the mandate for councilors elected in 2020 and the first year in which race reporting was mandatory in the dataset. Our sample is restricted to municipalities where both *white* and *nonwhite* employees were hired in the legislative sector.[3]

We define *nonwhites* as individuals classified as *black* (*preto*) or *mixed-race* (*pardo*). The other official classifications—*Indigenous* and *Asian* (*amarelo*)—account for less than 3% of legislative-sector workers and are excluded from the analysis, as societal perceptions toward these groups differ from those toward *pretos* and *pardos*. Legislative workers are identified using the legal nature (*natureza jurídica*) code 1066, which encompasses both legislative and administrative staff, including councilor aides, analysts, and operational personnel. Our outcome variable is the share of *nonwhite* legislative workers in each municipality, computed by aggregating the number of *nonwhite* and *white* hires and dividing the former by the total. In 2023, municipalities employed on average 57.7% *white* legislative workers.

To complement the hiring data, we incorporate information from the Brazilian Institute of Geography and Statistics (IBGE) on the race of the heads of municipal secretariats for health and education. These data are drawn from the 2018 and 2021 editions of the Municipality Profiles survey (*Perfil dos Municípios*), which reports demographic and organizational characteristics of Brazilian municipalities. Including these earlier years enables us to assess potential political shifts in leadership before the 2023 hiring outcomes. In this exercise, the outcome variable is binary, indicating whether the race of secretariat heads shifted from *white* to *nonwhite* between 2018 and 2021.

The treatment variable is constructed from candidate-level data provided by the Superior Electoral Court (TSE) for the 2020 municipal elections. This dataset contains sociodemographic information for each candidate, including self-reported race. The TSE also provides official photographs of all municipal council candidates, which we process using a facial recognition model — the Python package *DeepFace* — to infer each candidate's race.[4] This tool employs deep convolutional neural networks trained on large datasets with

[3] Race has been an optional RAIS field since 2004 and was often reported in the private sector. However, until 2023, all race entries for public employees were missing (*não identificado*).

[4] We use the *opencv* face-detection algorithm, which is the default option in *DeepFace*. Our main results are robust to employing *RetinaFace* (Deng, Guo, Zhou, Yu, Kotsia and Zafeiriou, 2019) as an alternative, with the sole exception of the findings concerning the heads of secretariats discussed below.

over 100,000 facial images to estimate the likelihood that an individual belongs to one of six broad racial groups: *asian*, *indian*, *black*, *white*, *middle Eastern*, and *latino/hispanic* (Karkkainen and Joo, 2021; Serengil and Ozpinar, 2024).[5] Instead of assigning a single category, the model produces a probability distribution across all six. We then estimate a probit model in which the binary dependent variable is self-reported race (*white* or *nonwhite*) and the explanatory variables are the predicted probabilities from *DeepFace*; the fitted values yield the predicted probability of being *white*. See Appendix A for further discussion.

Figure 1 shows the distribution of the predicted probability of being *white*. Panel A indicates that, as expected, most self-reported *nonwhite* candidates have low predicted probabilities: specifically, 75% have a predicted probability below 53.6%. Panel B likewise displays the expected pattern for self-reported *white* candidates, with 75% having a predicted probability above 46%. We use this 46% threshold to define *misaligned white* candidates — those in the bottom quartile of the predicted probability distribution among self-declared *white* candidates.

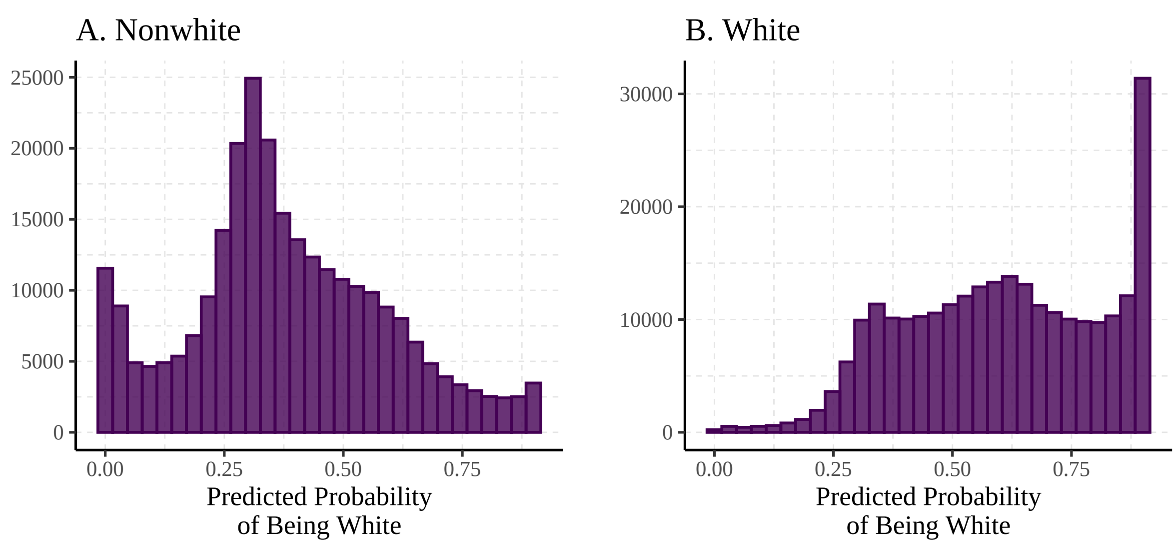

Figure 1: **Predicted Probability of Being *White*.** Distribution of the *predicted probability of being white* for candidates who self-reported as white and nonwhite. The *predicted probability of being white* is the fitted probabilities from a probit model where the individual's self-reported race is the binary dependent variable and the likelihood that an individual belongs to one of six broad ethnoracial group as explanatory variables.

[5]The effectiveness of the method may vary depending on factors such as image resolution, lighting, and the subject's facial orientation. In addition, the algorithm is based on broad, US-oriented racial classifications that do not fully capture Brazil's unique racial dynamics—most notably, the absence of a clear equivalent to the *pardo* category. Nevertheless, because candidate photographs are officially submitted to the Brazilian Electoral Court and must meet minimum quality standards, we expect the algorithm's outputs to be sufficiently reliable. Moreover, the use of broad racial categories aligns with our objective: we are not attempting to assign a definitive racial identity to each individual, but rather to measure the likelihood they are perceived as *white* given their physical appearance.

To construct the running variable for the regression discontinuity design along the lines of Novaes (2024), we use official election results to calculate the vote margin between the lowest-ranked elected candidate and the highest-ranked unelected candidate within each party coalition that secured at least one seat. The margin is positive if the *misaligned white* candidate won and negative if they lost. When multiple *misaligned white* candidates are present in a municipality, we select the one closest to the cutoff. If both winners and losers are present, we use the margin of the winning candidate; if all *misaligned white* candidates lost, we use the closest loser. This approach allows us to exploit plausibly exogenous variation in the election of *misaligned white* candidates.

Table 1 reports summary statistics for municipal council candidates, presented separately for the full sample and the restricted sample. In the full sample, the majority of candidates are male (about 65%), with an average age of 45. Roughly half are married, and around 23% have higher education. Nearly 47% self-identify as White, about 4% are incumbents seeking reelection, and 11% are affiliated with a left-wing party.[6] In the restricted sample—constructed based on the margin of victory as described above—we observe higher proportions of male, married, and higher-educated candidates, as well as a greater share of incumbents. Conversely, the share affiliated with left-wing parties is smaller. By construction, all candidates in the restricted sample self-identify as *white*. Importantly, consistent with our identification strategy, the characteristics in Table 1 show no discontinuous changes at the cutoff, as demonstrated in the results section below.

Table 1: **Summary Statistics: Full vs. Restricted Sample**

	Full Sample			Restricted Sample		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Male	487,431	0.653	0.476	3,542	0.872	0.334
Age	487,425	45.280	11.570	3,542	45.511	10.478
Married	487,431	0.503	0.500	3,542	0.621	0.485
Higher Educ.	487,431	0.228	0.419	3,542	0.276	0.447
White	487,431	0.473	0.499	3,542	1.000	0.000
Reelection	487,431	0.043	0.202	3,542	0.161	0.367
Leftwing	487,431	0.115	0.319	3,542	0.077	0.267

[6]We classify as left-wing the following parties: Workers' Party (PT, the main left-wing party), Socialism and Liberty Party (PSOL), Communist Party of Brazil (PCdoB), Brazilian Communist Party (PCB), Workers' Cause Party (PCO), Unified Socialist Workers' Party (PSTU), Green Party (PV), and Sustainability Network (REDE). The last two parties, PV and REDE, are primarily concerned with environmental causes.

3 Results

We begin the results section by verifying that candidate characteristics are balanced around the cutoff. In Figure 2A, we apply our regression discontinuity model to test for discontinuities in key characteristics: share of male candidates, average age, share married, share with higher education, share of incumbents seeking reelection, and share affiliated with left-wing parties. For all these variables, we detect no significant jumps at the cutoff, indicating that our sample construction does not systematically alter the composition of candidates.

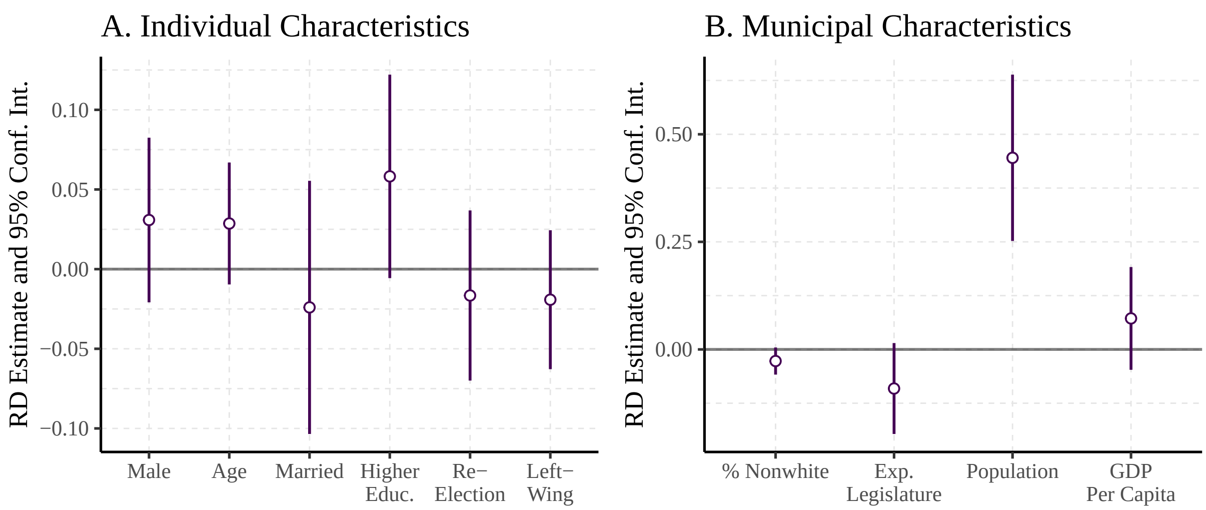

Figure 2: **Balance Tests.** In Panel A, each point represents the estimated effect of electing one additional *misaligned white* councilor on the following characteristics of elected councilors: the share who are male, average age, share who are married, have higher education, are incumbents seeking reelection, and are affiliated with left-wing parties. In Panel B, each point represents the estimated impact of electing one additional *misaligned white* councilor on the following outcomes, respectively: the share of the *nonwhite* population, legislative sector expenditures, total population size, and GDP.

Figure 2B examines potential discontinuities in broader municipal demographic characteristics. First, we test for changes in the share of *nonwhite* residents and find no significant discontinuities, ruling out population composition as a confounding factor. We also assess discontinuities in the logarithms of per capita legislative spending, per capita GDP, and total population. While we find no significant jumps in legislative spending or GDP, we observe a discontinuity in total population: municipalities to the right of the cutoff have populations approximately 40% larger, indicating an imbalance in our sample. Although this poses a potential threat to identification, we demonstrate that our results remain robust when controlling for population size.

We also perform a manipulation density test following Cattaneo, Jansson and Ma (2020) to examine whether the density of the running variable exhibits a discontinuity at the cutoff. The density around the cutoff, along with local polynomial estimates, is shown in Figure B.1 in Appendix B. As expected, we observe a spike just above the threshold. This discontinuity is a mechanical artifact of how we construct the running variable. Specifically, the running variable is the vote margin between the lowest-ranked elected candidate and the highest-ranked unelected candidate within each party coalition that won at least one seat. The margin is positive if the *misaligned white* candidate won and negative if they lost. When multiple *misaligned white* candidates appear in a municipality, we select the one closest to the cutoff — prioritizing winning candidates whenever both winners and losers are present. If all *misaligned white* candidates lost, we select the closest loser. This method isolates plausibly exogenous variation in the election of misreporting candidates but creates an asymmetry in the data: by construction, winners are more likely to be selected just above the cutoff, which mechanically increases the density on the right side of the threshold.

To address this, we re-estimate the manipulation test after excluding observations within a "donut" region of ±0.005 around the cutoff. Once this region is removed, the test statistic equals 0.9, preventing us from rejecting the null hypothesis of no manipulation. We thus conclude that the initial jump is localized and not due to strategic manipulation. Importantly, our main results (discussed below) remain robust when excluding this "donut." In fact, the estimated effect increases in magnitude and remains statistically significant at the 95% confidence level (coefficient of –16 percentage points, standard error of 7 percentage points), further reinforcing the credibility of our research design. Below, we present the baseline estimated effect of electing one additional *misaligned white* candidate on racial representation in public employment outcomes.

Figure 3 presents our main findings. Panel A reveals a clear discontinuity in the hiring of *nonwhite* employees within the legislative sector on the right side of the cutoff, corresponding to an estimated decrease of 8.4 percentage points. Panel B replicates the analysis for the private sector, where no such discontinuity is found. These results indicate a significant decline in the share of *nonwhite* hires in the municipal legislative sector following the narrow election of a *misaligned white* candidate. The magnitude of this effect is substantial — approximately a 20% reduction relative to the average share of *nonwhite* hires. Notably, this effect appears confined to the public sector, suggesting it likely arises from the discretionary hiring authority of elected officials.

Next, we investigate whether the decline in hiring is linked to discretionary employment practices by

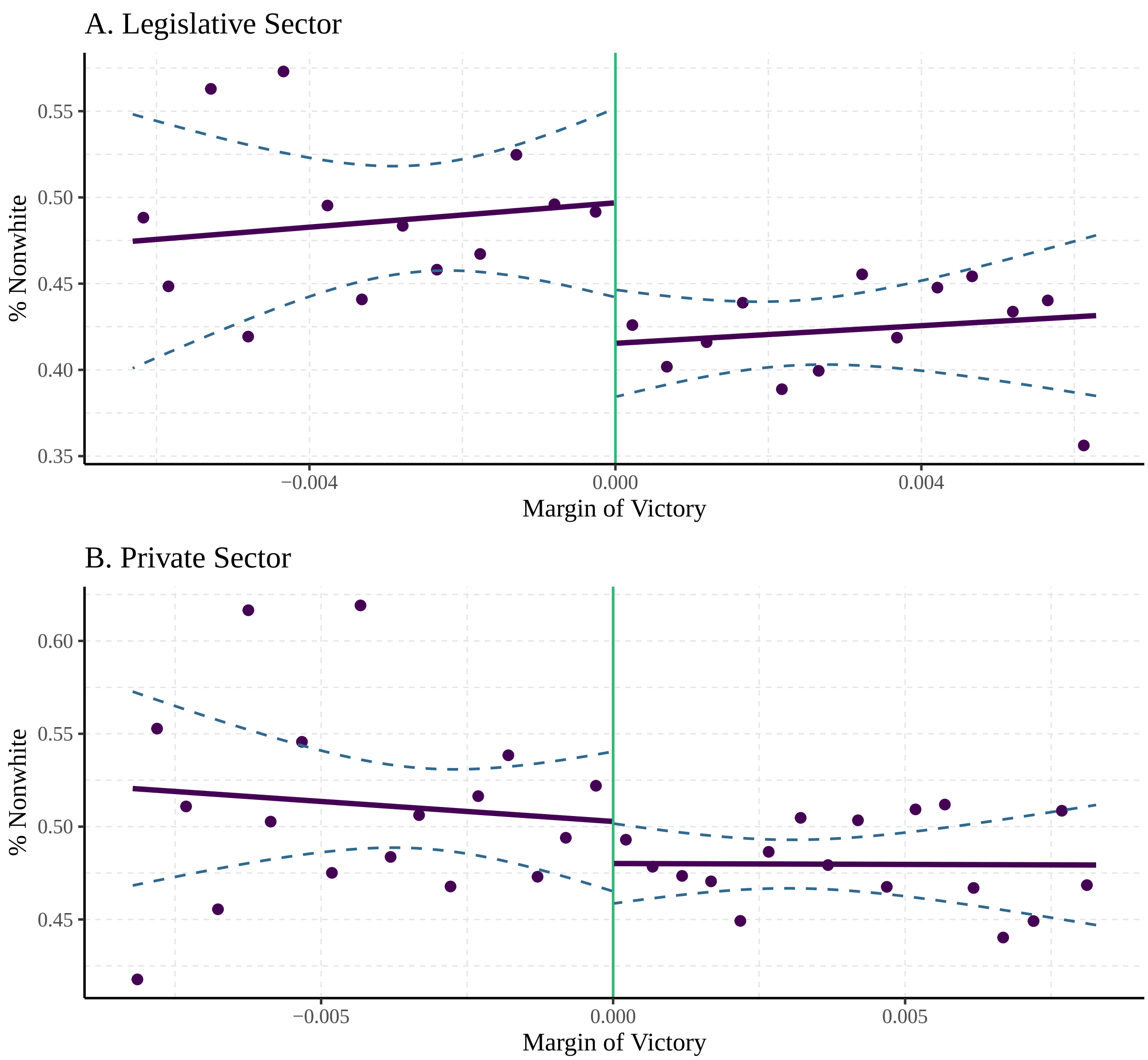

Figure 3: **Margin of Victory Discontinuity.** This figure presents a function of margin of victory size using the MSE-optimal bandwidth approach with a uniform kernel Calonico et al. (2014), with a vertical line marking the cutoff at 0. A solid line is fitted separately on each side of the threshold, and the dashed lines show the 95% confidence interval. Scatter plots display averages within 0.05 percentage point intervals. Panels A and B present results for percentage of *nonwhite* workers hired in the public (legislative) and private sector, respectively.

splitting the sample into permanent and temporary workers. Permanent workers are civil servants who obtain tenure and job stability after passing a competitive exam, making their hiring less susceptible to political influence. Figure 4 displays box plots of the estimated effects, revealing that the negative impact is concentrated among temporary employees — consistent with expectations, as their hiring is more easily influenced by elected officials. Furthermore, the effect is stronger for temporary employees in managerial roles — positions involving greater discretion — with an estimated reduction of 25%, compared to a 20% decrease among non-managerial temporary workers.

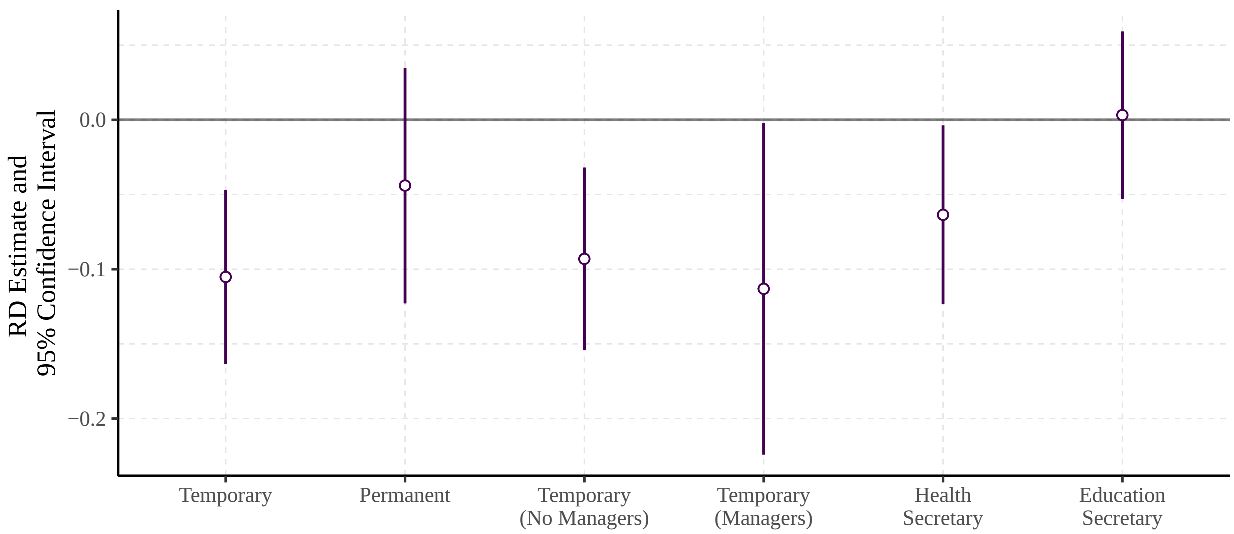

Figure 4: **RD Estimates.** Each point shows the estimated effect of electing an extra *misaligned white* councilor on the percentage of *nonwhite* workers hired in the legislative sector, distinguishing between temporary, permanent, managers and heads of secretariats. In the latter case, the effect is on the likelihood that a municipality experienced an increase in the number of *nonwhite* managers in the health or education secretariats between 2018 and 2021.

Finally, we use data from IBGE's Municipality Profiles survey to assess whether the election of *misaligned white* councilors influenced a change in the racial composition of leadership positions in municipal public administration. We focus on changes in the heads of the municipal health and education secretariats between 2018 and 2021. Figure 4 shows that electing a *misaligned white* candidate is associated with a roughly 20% reduction in the likelihood that the race of the head of the health secretariat shifted from *white* to *nonwhite*. No comparable effect is found for the education secretariat. These results suggest that *misaligned white* candidates may shape not only legislative hiring but also high-level bureaucratic appointments, particularly in politically salient domains such as health policy. It is important to note that these results are sensitive to the choice of face-detection algorithm. When *RetinaFace* is used, the effect on the head of the health secretariat is smaller and statistically insignificant, while the effect on the head of the education secretariat becomes negative and is statistically significant at the 5% level. These findings should therefore be interpreted with caution. The full set of results obtained with *RetinaFace* is available upon request.

We estimate alternative model specifications to test the robustness of our results and to account for the discontinuity in population size, which could potentially bias our estimates. Figure 5 shows that all tested models — including those with a second-order polynomial and a triangular kernel — yield similar results. Moreover, adding controls for population size, share of the *nonwhite* population, GDP per capita, and legislative spending increases the magnitude of the estimated effects, indicating that these factors are not driving our findings.

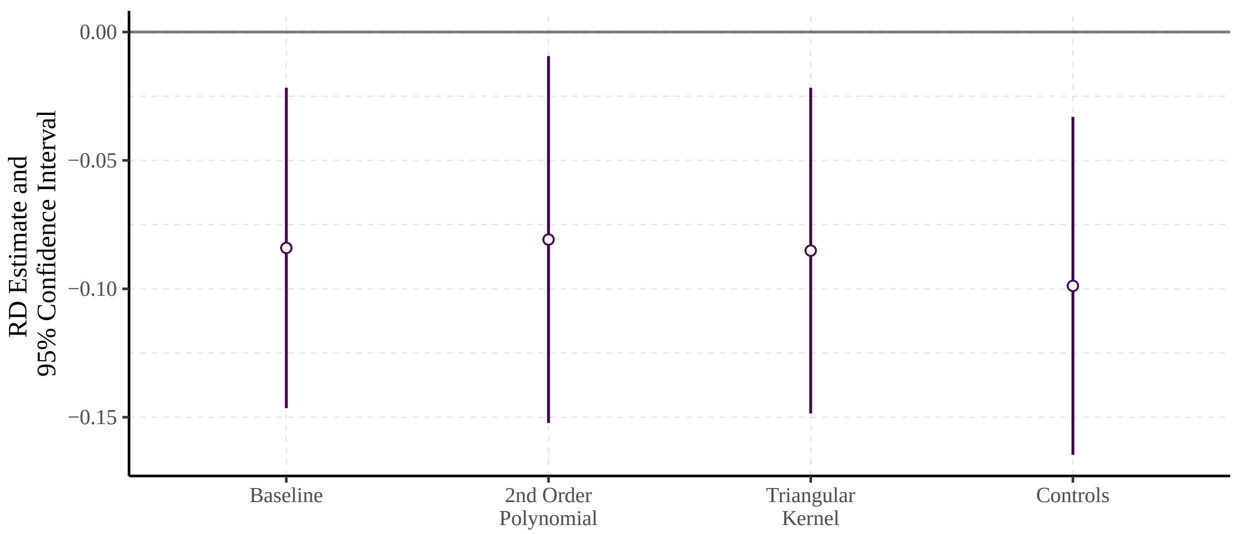

Figure 5: **Alternative Specifications.** Each point represents the estimated impact of electing one additional *misaligned white* councilor on the share of *nonwhite* employees in the legislative sector. We report results from four different model specifications: (i) a baseline model, (ii) a model using a second-order polynomial of the running variable, (iii) a model with a triangular kernel, and (iv) a model that adds controls for the share of the *nonwhite* population, total population size, per capita legislative expenditures, and per capita GDP.

We also examine whether the wage or hours-worked gap between *nonwhite* and *white* employees changes following the election of a *misaligned white* candidate, finding no evidence of such effects in either the legislative or private sectors. These null results suggest that the impact operates on the extensive margin — affecting who is hired — rather than on the intensive margin of job quality or compensation. As an additional mechanism test, we repeat the analysis using candidates who self-identified as *nonwhite* but were classified by the facial recognition model as having a high probability of being *white*. In this case, we again find no evidence of discontinuities in public hiring by race, reinforcing the interpretation that the observed effects are specifically driven by *misaligned whiteness*.

Taken together, the results show that racial identity — particularly when misaligned with social perceptions — can shape patterns of racial representation in the Brazilian public sector through electoral channels. The effects are substantial, concentrated in both the legislative sector and upper-level appointments within the health administration. In the legislative sector, the impact is driven mainly by temporary positions, where hiring is more discretionary, as opposed to permanent posts that grant tenure and stability. We find no significant effects on the education secretariat, broader labor market outcomes, or demographic composition, reinforcing the interpretation that these changes operate primarily through political and bureaucratic pathways.

4 Conclusion

This paper provides evidence that the election of misaligned *white* candidates — those who self-identify as *white* but are unlikely to be classified as *white* by facial recognition — has measurable consequences for racial representation in Brazil's public sector. Using a regression discontinuity design based on narrow electoral victories, we find that the election of such candidates leads to a substantial decline in the hiring of *nonwhite* employees in the municipal legislative sector, amounting to roughly a 20% reduction relative to the average share. This effect is concentrated in temporary positions, especially in managerial roles where hiring discretion is greatest, and extends to upper-level appointments in municipal health secretariats. We find no comparable effects in the private sector, in permanent public positions, suggesting that the mechanism operates through discretionary hiring and political influence rather than through broader labor market channels.

The robustness of these findings is supported by extensive balance tests, alternative model specifications, and the exclusion of observations near the cutoff to rule out mechanical artifacts. The absence of significant effects on wages, hours worked, or demographic shifts further reinforces that the impact occurs at the extensive margin—determining who is hired—rather than altering the quality of employment. Taken together, the results highlight how racial identity, particularly when mismatched, can shape bureaucratic and political outcomes. By showing that such effects are localized to domains of high political discretion, our findings underscore the importance of considering both racial identity and the institutional contexts in which electoral outcomes translate into concrete shifts in representation.

Online Appendix to

Racial Self-Classification, Group Consciousness, and Public Employment Representation

Diogo Baerlocher
University of South Florida

Rodrigo Schneider
Skidmore College

Appendix A Probability of Being White

We define the probability of being white as a combination of the probabilities that the face recognition algorithm classifies the individuals in one the following race groups: *white*, *black*, Asian, Latino/Hispanic, Indian and Middle Eastern. Denote these probabilities for individual k as p_{jk} with j$\in J := w, b, a, l, i, m$ representing each possible race, respectively. The predicted probability of being white is then

$$P(white)_k = F\left(\sum_{j \in J} \omega_j p_{jk}\right),$$

where $F : \mathbb{R} \rightarrow [0, 1]$ is a function and ω_j are weights assigned to each race's probability. One way to determine weights is to use *ad hoc* assumptions about individual self-classification in Brazil. For example, it is likely that individuals with Middle Eastern or Asian appearance will self-report as *white*. We can thus set $F(x) = x$ for $x \in [0, 1]$ and $\omega_j = 1$ if $j \in w, a, m$, and $\omega_j = 0$ otherwise. Alternatively, we can be more conservative and set $\omega_j = 1$ if $j = w$, and $\omega_j = 0$ otherwise.

One data-driven approach would set weights from a probit regression where the dependent variable is binary indicating if the individual self-reported as white and the regressors are p_{jk}, excluding the constant to avoid collinearity. Table A.1 shows the probit model estimation representing ω_j. In this case, F is the standard normal cumulative distribution function. As expected, the probability of being *white* is largely determined by larger probabilities of being identified as *white* by the algorithm. Middle Eastern is the only other ethnoracial group with positive weight. Large probabilities of being classified as *black* or Indian strongly reduces the probability of being *white*, which is again expected given the similar skin tone among these two groups.

Table A.1: **Probit Regression - Dependent variable** = 1 **if self-reported as *white***

	Coefficient	Standard Error	z	$P > \|z\|$	95% Conf. Interval	
white	1.3504	0.0063	215.85	0.00	1.3381	1.3627
black	-2.2498	0.0198	-113.6	0.00	-2.2886	-2.2110
Asian	-0.4470	0.0108	-41.27	0.00	-0.4683	-0.4258
Middle Eastern	0.3745	0.0095	39.64	0.00	0.3560	0.3930
Indian	-2.2146	0.0285	-77.81	0.00	-2.2704	-2.1588
Latino/Hispanic	-0.4486	0.0068	-66.22	0.00	-0.4619	-0.4353

Probit regression with 529,780 observations. The dependent variable indicates if the individual self-reported as *white*. The regressors are probability of being identified in one of the ethnoracial groups listed in column 1 by the facial recognition algorithm. The constant is remove since the probability must sum one.

Appendix B Manipulation Density Test

To assess potential manipulation of the running variable around the cutoff, we implement a manipulation density test using the procedure of Cattaneo et al. (2020). The test reveals a discontinuity in the density at the threshold, suggesting possible sorting or manipulation. We also present the corresponding density plot in Figure B.1, which includes binned estimates of the running variable's distribution and local polynomial fits with valid confidence bands. To investigate whether the discontinuity is localized, we re-estimate the test after excluding observations within ±0.005 of the cutoff. Once this "donut" region is removed, the manipulation test statistic equals 0.9, and we are unable to reject the null hypothesis of no manipulation. We conclude that the initial jump is localized and not driven by strategic sorting around the threshold.

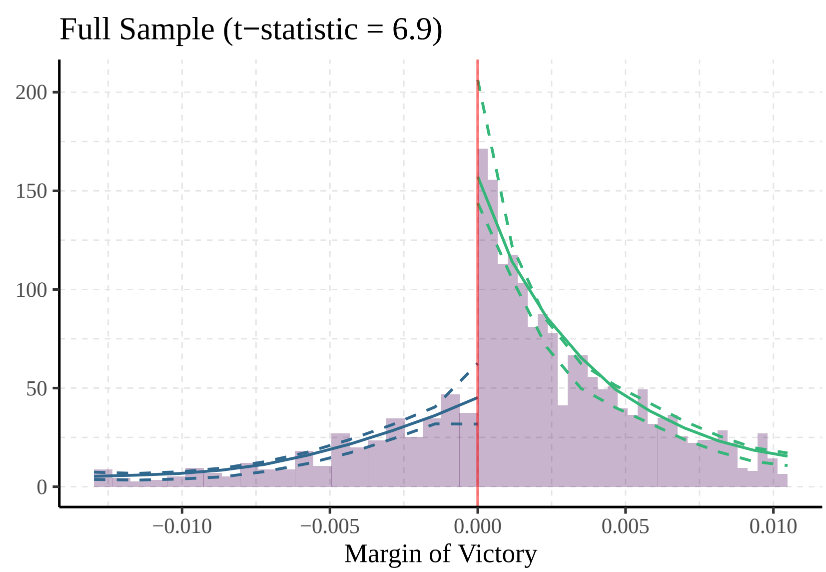

Figure B.1: **Margin of Victory Density.** Density of margin of victory around the cutoff zero with bandwidth selected using the uniform kernel function to estimate local polynomial estimators, represented by solid lines. Dashed lines represent 95% confidence intervals computed using jackknife standard errors.

References

Alesina, Alberto, Reza Baqir, and William Easterly, "Public goods and ethnic divisions," *The Quarterly journal of economics*, 1999, *114* (4), 1243–1284.

Beach, Brian and Daniel B Jones, "Gridlock: Ethnic diversity in government and the provision of public goods," *American Economic Journal: Economic Policy*, 2017, *9* (1), 112–136.

_ , _ , Tate Twinam, and Randall Walsh, "Racial and ethnic representation in local government," *American Economic Journal: Economic Policy*, 2024, *16* (2), 1–36.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust data-driven inference in the regression-discontinuity design," *The Stata Journal*, 2014, *14* (4), 909–946.

Cascio, Elizabeth U and Ebonya Washington, "Valuing the vote: The redistribution of voting rights and state funds following the voting rights act of 1965," *The Quarterly Journal of Economics*, 2014, *129* (1), 379–433.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso, "Patronage and selection in public sector organizations," *American Economic Review*, 2020, *110* (10), 3071–3099.

Cornwell, Christopher, Jason Rivera, and Ian M Schmutte, "Wage discrimination when identity is subjective: Evidence from changes in employer-reported race," *Journal of Human Resources*, 2017, *52* (3), 719–755.

de Lucena Coelho, Thiago, Fernanda Estevan, Marcos Nakaguma, and Alexandre Rabelo, "Do Black Politicians Matter? Political Leadership and Racial Composition in Top Public Sector Positions," 2024. Available at SSRN: https://ssrn.com/abstract=5183898.

de Paiva Rio Camargo, Alexandre, "Mensuração racial e campo estatístico nos censos brasileiros (1872-1940): uma abordagem convergente," *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 2009, *4*, 361–385.

Deng, Jiankang, Jia Guo, Yuxiang Zhou, Jinke Yu, Irene Kotsia, and Stefanos Zafeiriou, "RetinaFace: Single-stage Dense Face Localisation in the Wild," 2019. Available at https://arxiv.org/abs/1905.00641.

Ikawa, Jorge, Clarice Martins, Pedro C Sant'Anna, and Rogerio Santarrosa, "Elections that Inspire: Effects of Black Mayors on Educational Attainment," 2024. Available https://pedrosant-anna.github.io/files/BlackMayors_May9.pdf.

Janusz, Andrew, "Electoral incentives and elite racial identification: Why Brazilian politicians change their race," *Electoral Studies*, 2021, *72*, 102340.

—, "Race and resources in brazilian mayoral elections," *Political Research Quarterly*, 2022, *75* (3), 846–859.

Karkkainen, Kimmo and Jungseock Joo, "Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation," in "Proceedings of the IEEE/CVF winter conference on applications of computer vision" 2021, pp. 1548–1558.

Novaes, Lucas M, "The violence of Law-and-Order politics: the case of law enforcement candidates in Brazil," *American Political Science Review*, 2024, *118* (1), 1–20.

Nye, John VC, Ilia Rainer, and Thomas Stratmann, "Do black mayors improve black relative to white employment outcomes? Evidence from large US cities," *The Journal of Law, Economics, & Organization*, 2015, *31* (2), 383–430.

Pinhoni, Marina and Gabriel Croquer, "Censo 2022: Pela 1a vez, Brasil se declara mais pardo que branco; populações preta e indígena também crescem," *G1*, 2023, *December 22.* Available at https://g1.globo.com/economia/censo/noticia/2023/12/22/censo-2022-cor-ou-raca.ghtml.

Romero, Luiz, "Why did thousands of Brazilian politicians change their race?," *CNN*, 2021, *April 10.* Available at https://www.cnn.com/2021/04/09/americas/brazil-race-politics-change-intl-latam/index.html.

Schneider, Rodrigo and Henrique Veras, "Do bigger legislatures lead to bigger government? Evidence from a Brazilian municipal council reform," *Economía*, 2023, *22* (1), 117–134.

Sen, Maya and Omar Wasow, “Race as a bundle of sticks: Designs that estimate effects of seemingly immutable characteristics,” *Annual Review of Political Science*, 2016, *19* (1), 499–522.

Serengil, Sefik and Alper Ozpinar, “A Benchmark of Facial Recognition Pipelines and Co-Usability Performances of Modules,” *Journal of Information Technologies*, 2024, *17* (2), 95–107.

Skidmore, Thomas E, *Brazil: Five centuries of change*, Oxford University Press, 2009.

Smith, Jasmine Carrera, Jared Clemons, Arvind Krishnamurthy, Miguel Martinez, Leann McLaren, and Ismail K White, “Willing but unable: Reassessing the relationship between racial group consciousness and Black political participation,” *American Political Science Review*, 2024, *118* (4), 1671–1687.